RC23464 (W0412-032) December 13, 2004
Mathematics

IBM Research Report

About the Optimal Density Associated to the Chiral Index of a Sample from a Bivariate Distribution

Don Coppersmith
IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598

Michel Petitjean
ITODYS
(CNRS, UMR 7086, Université Paris 7)
1 rue Guy de la Brosse
75005 Paris
France

Probability / Statistics

About the Optimal Density Associated to the Chiral Index of a Sample from a Bivariate Distribution

A propos de la densité optimale associée à l'indice chiral d'un échantillon d'une distribution bivariée

Don Coppersmith[1] and Michel Petitjean[2,*]

[1] IBM TJ Watson Research Center, Yorktown Heights, New York 10598, USA.

Email: dcopper@us.ibm.com

http://www.research.ibm.com/people/c/coppersmith/

[2] ITODYS (CNRS, UMR 7086, Université Paris 7), 1 rue Guy de la Brosse, 75005 Paris, France. Phone: 33 (0)1 4427 4857. Fax: 33 (0)1 4427 6814

petitjean@itodys.jussieu.fr

http://petitjeanmichel.free.fr/itoweb.petitjean.html

[*] To whom correspondence should be addressed.

Abstract

The complex quadratic form $V_z(P) = z'Pz$, where z is a fixed vector in $\mathbf{C}^n$ and z' is its transpose, and P is any permutation matrix, is shown to be a convex combination of the quadratic forms $V_z(P_\sigma)$, where P_σ denotes the symmetric permutation matrices. We deduce that the optimal probability density associated to the chiral index of a sample from a bivariate distribution is symmetric. This result is used to locate the upper bound of the chiral index of any bivariate distribution in the interval $[1-1/\pi, 1-1/2\pi]$.

Résumé

Nous montrons que la forme quadratique complexe $V_z(P) = z'Pz$, où z est un vecteur donné dans $\mathbf{C}^n$ et z' est son transposé, et P est une matrice de permutation, est une combinaison convexe des formes quadratiques $V_z(P_\sigma)$, où les P_σ sont des matrices de permutation symétriques. On en déduit que la densité de probabilité optimale associée à l'indice chiral d'un échantillon d'une distribution bivariée est symétrique. Ce résultat est utilisé pour localiser la borne supérieure de l'indice chiral d'une distribution bivariée quelconque dans l'intervalle $[1 - 1/\pi, 1 - 1/2\pi]$.

AMS 2000 subject classification: 60E15

1. Introduction

The chiral index χ of a finite variance d-variate probability distribution $\mathcal{P}$ is the Wasserstein distance between the distribution $\mathcal{P}$ and its inverted image $\bar{\mathcal{P}}$, minimized for all rotations and translations of $\bar{\mathcal{P}}$, and normalized to the inertia of $\mathcal{P}$ [1]. It takes values over $[0, 1]$. It is a skewness measure offering various applications in computer sciences [2]. In the case of a sample of size n, the optimal joint density between and $\mathcal{P}$ and $\bar{\mathcal{P}}$ is known to exist [3]. The matrix associated to this optimal density is shown to be $(1/n)$ times a permutation matrix [2,4]. In the univariate case, this permutation matrix is symmetric [5]. We extend the result in this paper to the bivariate case: the optimal joint density is symmetric. The upper bound of the chiral index of a d-variate distribution is unknown, except in the univariate case, for which it is $1/2$ [1]. In the bivariate case, the symmetry of the optimal joint density of a sample is used to locate the upper bound in $[1 - 1/\pi, 1 - 1/2\pi]$.

2. Symmetry of the permutation

We first need to establish two theorems in the complex plane.

Fix a complex n-vector $z = (z_1, z_2, \ldots, z_n) \in \mathbf{C}^n$. Given a permutation σ on n indices $\{1, 2, \ldots, n\}$, define the function V_z of σ as $V_z(\sigma) = \sum_{j=1}^{n} z_j z_{\sigma(j)}$. Notice that we are not taking complex conjugates. When z is understood, we will write $V(\sigma)$ for $V_z(\sigma)$.

A *symmetric* permutation σ is one satisfying $\sigma = \sigma^{-1}$. It is the product of disjoint 1-cycles and 2-cycles. The permutation matrix P_σ is symmetric and $V(P_\sigma) = V(P'_\sigma)$.

Theorem 1. For any permutation τ, $V(\tau)$ is a convex combination of $\{V(\sigma) : \sigma = \sigma^{-1}\}$.

This following lemma will be crucial:

Lemma 1. If τ is an n-cycle, then there is a symmetric permutation σ satisfying $\Re(V(\sigma)) \leq \Re(V(\tau))$.

To prove the lemma: Let $\tau = (1, 2, 3, \ldots, n)$. When $n \in \{1, 2\}$ the result is immediate: τ itself is a symmetric permutation. When n is even, define two permutations (as products of $\frac{n}{2}$ disjoint 2-cycles)

$$\begin{aligned} \alpha &= (1,2)(3,4)\cdots(n-1,n) \\ \beta &= (2,3)(4,5)\cdots(n,1) \end{aligned}$$

and compute that

$$V(\tau) = \frac{1}{2}\left[V(\alpha) + V(\beta)\right].$$

So $V(\tau)$ is a convex combination of $V(\alpha)$ and $V(\beta)$, whence $\min\{\Re(V(\alpha)), \Re(V(\beta))\} \leq \Re(V(\tau))$.

We are left with the case where n is odd, $n \geq 3$.

Consider the following $2n$ permutations:

$$\left.\begin{aligned} \alpha_j &= (j)(j+1, j+2)(j+3, j+4)\cdots(j-2, j-1) \\ \beta_j &= (j-1, j+1)(j)(j+2, j+3)(j+4, j+5)\cdots(j-3, j-2) \end{aligned}\right\} 1 \leq j \leq n$$

We compute:

$$V(\alpha_j) + V(\alpha_{j+1}) - 2V(\tau) = z_j^2 + z_{j+1}^2 - 2z_j z_{j+1} = (z_{j+1} - z_j)^2,$$

where we are considering the indices modulo n, so that if $j = n$ then $z_{j+1} = z_1$.

$$V(\alpha_j) + V(\beta_j) - 2V(\tau) = 2z_j^2 + 2z_{j-1}z_{j+1} - 2z_{j-1}z_j - 2z_j z_{j+1} = -2(z_{j+1} - z_j)(z_j - z_{j-1}).$$

Now suppose the lemma is false, so that for all j, $\Re(V(\alpha_j)) > \Re(V(\tau))$ and $\Re(V(\beta_j)) > \Re(V(\tau))$. Then for all j,

$$\Re[(z_{j+1} - z_j)^2] > 0.$$

In particular

$$\Re(z_{j+1} - z_j) \neq 0.$$

Also,

$$\Re[-2(z_{j+1} - z_j)(z_j - z_{j-1})] > 0.$$

Fix j. Define

$$\begin{aligned} b &= z_{j+1} - z_j \\ c &= z_j - z_{j-1}. \end{aligned}$$

We have just seen that $\Re(b^2) > 0$ and $\Re(c^2) > 0$ and $\Re(bc) < 0$. Observe also that $(\Re(b)c - \Re(c)b)$ is pure imaginary, so its square is real and nonpositive:

$$\Re(b)^2 c^2 + \Re(c)^2 b^2 - 2\Re(b)\Re(c)bc \leq 0.$$

Taking the real parts of all terms and rearranging,

$$2\Re(b)\Re(c)\Re(bc) \geq \Re(b)^2\Re(c^2) + \Re(c)^2\Re(b^2) > 0,$$

and from $\Re(bc) < 0$ we conclude

$$\Re(b)\Re(c) < 0.$$

So the sign of $\Re(b) = \Re(z_{j+1} - z_j)$ and the sign of $\Re(c) = \Re(z_j - z_{j-1})$ are opposite. As j cycles around $1, 2, \ldots, n, 1$, the signs of $\Re(z_{j+1} - z_j)$ alternate. But n is odd, so this alternation is impossible. The contradiction proves the lemma.

Lemma 2. If τ is an n-cycle, then $V(\tau)$ is a convex combination of $\{V(\sigma) : \sigma = \sigma^{-1}\}$.

Proof: Suppose the conclusion is false. Then there is a line ℓ through $V(\tau)$ in the complex plane, with all $\{V(\sigma) : \sigma = \sigma^{-1}\}$ lying on one side of the line. If ℓ has direction θ, we have $\Re[V_z(\sigma)e^{i(\pi/2-\theta)}] > \Re[V_z(\tau)e^{i(\pi/2-\theta)}]$ for all σ with $\sigma = \sigma^{-1}$. Now set $w = ze^{i(\pi/2-\theta)/2}$ so that $V_w(\tau) = e^{i(\pi/2-\theta)}V_z(\tau)$, and apply Lemma 1.

Proof of Theorem 1: To see that the theorem follows from Lemma 2, express an arbitrary permutation τ as a product of disjoint cycles τ_j; apply the lemma to each cycle (and its reduced set of variables $\{z_k\}$); and use additivity of $V(\sigma)$.

Then Theorem 2 is deduced immediately from Theorem 1:

Theorem 2. The modulus of $z'Pz$ is maximized by a symmetric permutation matrix P_σ.

It is pointed out that the optimal permutation may be not unique. An example is $z' = (1, \mathbf{i}, 2 - \mathbf{i})$, where $\mathbf{i} = \sqrt{-1}$. The identity permutation ι and the permutation τ interchanging the last two elements share $|V(\iota)| = |V(\tau)| = 5 > V(\pi)$ for all other permutations π. Non symmetric permutations may be optimal (e.g. when z has several identical elements).

3. Application to the chiral index

We set $z = x + \mathbf{i}y$, x and y being fixed vectors in $\mathbf{R}^d$. Then, the modulus of the complex quadratic form $z'Pz$ is the difference of the eigenvalues λ_1 and λ_2 (sorted in decreasing order), of the matrix $X'(\frac{P+P'}{2})X$, where the matrix $X = [x|y]$ has two columns and n lines. Thus:

$$|z'Pz| = \lambda_1 - \lambda_2$$

The chiral index χ is computed at null expectation from equations (6) and (7) in [5]. For a sample of a bivariate distribution $\mathcal{P}$ with inertia $T = \|z\|^2$ and $\|z\|^2 = x'x + y'y$, we have:

$$nT(1-\chi) = \max_{\{P\}}(\lambda_1 - \lambda_2)$$

The matrix associated to the joint density between $\mathcal{P}$ and $\bar{\mathcal{P}}$ is $[P/n]$, and the Wasserstein distance between $\mathcal{P}$ and $\bar{\mathcal{P}}$ is $\max_{\{P/n\}}(\lambda_1 - \lambda_2)$. It follows from Theorem 2 that the optimal joint density matrix $[P/n]$ is symmetric.

We consider now the more general situation where the n points are partitioned into groups of colors [1,4,5]. Permutations involving cycles over two groups are no more considered, and the optimal permutation is taken over a subset of the $n!$ permutations. Obviously, Theorems 1 and 2 stand again, and the optimal joint density matrix is still symmetric. Colors are not further considered in this paper.

4. Localization of the upper bound of the chiral index: part 1

We exhibit here a family of centered sets for which the ratio $\max_{\{P\}} |z'Pz|/\|z\|^2$ is arbitrarily close to $1/\pi$. "Centering" means working at null expectation. It means here that $\mathbf{1}'z = 0$, where $\mathbf{1}$ is a vector in $\mathbf{C}^n$ each of whose elements is 1. It is also recalled that the ratio is insensitive to an arbitrary planar rotation (phase).

Fix $\varepsilon > 0$. Choose an even integer $m > 1/\varepsilon$. Let $\omega = e^{\mathbf{i}\frac{2\pi}{2m}}$ be a complex root of unity, so that $\omega^{2m} = 1$. Select an integer $r > m^4/\varepsilon^2$ and an even integer $k > r^{m-1}/\varepsilon$. The complex vector z has $n = (1 + r + r^2 + \cdots + r^{m-1} + 2k)$ elements as follows. There are $m+3$ blocks labelled $j = 0, \ldots, m+2$, each consisting of identical elements. For $j < m$, block j has r^j identical elements with value $\omega^j/r^{j/2}$. Let S denote the sum of these elements: $S = \sum_{j=0}^{j=m-1} \omega^j r^{j/2}$. Block m contains k identical elements with value $-S/k$; block $m+1$ contains $k/2$ elements with value iS/k; and block $m+2$ contains $k/2$ elements with value $-iS/k$. The sum of elements of z is zero: block m cancels the first m blocks, and blocks $m+1$, $m+2$ cancel each other. Also, the sum of squares of elements of z is zero: the squares of elements in the first m blocks add to $\sum_{j=0}^{m-1} \omega^{2j} = 0$, while blocks $m+1$ and $m+2$ cancel block m. One can compute $x'x = y'y = m/2 + O(\varepsilon)$ and $x'y = 0$.

We know from Theorem 2 that the optimal permutation P pairs the elements of z, some being paired with themselves when P contains 1-cycles. Let B_j be the number of elements paired within the block j. We set $\beta_j = B_j/r^j$, so that $0 \le \beta_j \le 1$ for $j = 0 \ldots m-1$. The contribution of these elements to $z'Pz$ is $\beta_j \omega^{2j}$.

One can see that the contribution to $z'Pz$ of the elements paired between two different blocks j_1 and j_2 is $O(1/r^{(1/2)|j_1 - j_2|}) = O(\varepsilon/m^2)$ when $j_1 < m$ and $j_2 < m$, so that the $m^2 - m$ off-diagonal blocks contribute a total of $O(\varepsilon)$ to $z'Pz$. The contribution of the elements paired between the blocks $j < m$ and blocks $m, m+1, m+2$ is $O\left(r^{m-1}\left(\frac{1}{\sqrt{r^{m-1}}}\right)\frac{|S|}{k}\right) = O(\varepsilon)$. The contribution of the elements paired within the last three blocks is $O\left(k\left(\frac{|S|}{k}\right)^2\right) = O(\varepsilon)$. All these contributions sum to at most $O(\varepsilon)$, except for the diagonal terms $j_1 = j_2 < m$.

We are looking for the limit when m tends to infinity, of the ratio $\max_{\{P\}} |z'Pz|/\|z\|^2$.

When m is arbitrarily large, we look for the m values β_j maximizing:

$$\frac{|\sum_{j=0}^{j=m-1} \beta_j \omega^{2j} + O(\varepsilon)|}{m + O(\varepsilon)}$$

The complex number $\gamma = \sum_{j=0}^{j=m-1} \beta_j \omega^{2j}$ is the sum of m terms having all modulus in the interval $[0,1]$. Neglecting the term $O(\varepsilon)$, we can see that only the terms offering a difference of phase ϕ such that $\cos(\phi) \geq \beta_j/2|\gamma|$ will contribute to the modulus of γ. Since $|\gamma|$ tends to infinity when m tends to infinity, only the terms having a difference of phase within $[-\pi/2, +\pi/2]$ with γ will contribute to the modulus of γ.

For these latter we set $\beta_j = 1$, and we set $\beta_j = 0$ elsewhere. Working with a free arbitrary phase, we have:

$$\gamma = 1 + \omega^2 + \omega^4 + \cdots + \omega^{2(m-1)/2} = \frac{1 - \omega^m}{1 - \omega^2} = \frac{2}{(-\mathbf{i}\omega)(2\sin(2\pi/2m))}$$

Its modulus is

$$|\gamma| = \frac{1}{\sin(\pi/m)} = \frac{m}{\pi} + O\left(\frac{1}{m}\right) = \frac{m}{\pi} + O(\varepsilon)$$

Therefore, the limit is:

$$\lim_{m \to \infty} \{\max_{\{P\}} |z'Pz|/\|z\|^2\} = 1/\pi$$

It means that our family of centered sets z has a chiral index arbitrarily close to $1 - 1/\pi$. Thus we have proved Lemma 3:

Lemma 3. The upper bound of the chiral index of a bivariate sample cannot be smaller than $1 - 1/\pi$.

5. Localization of the upper bound of the chiral index: part 2

We first show that no set can have the ratio $\max_{\{P\}} |z'Pz|/\|z\|^2$ smaller than $1/\pi$ under the additional condition that at least half of the n elements z_j are null. The centering condition is not set here.

We consider an arbitrary phase θ and its associated permutation P_θ such that z_j is paired with itself when $\Re(z_j^2 e^{\mathbf{i}\theta}) > 0$ and z_j is paired with a null element when $\Re(z_j^2 e^{\mathbf{i}\theta}) \leq 0$. Setting $z_j^2 = r_j e^{\mathbf{i}\phi_\mathrm{j}}$, we have $e^{\mathbf{i}\theta} z' P_\theta z = \sum r_j e^{\mathbf{i}(\theta+\phi_\mathrm{j})}$, and since $|z'P_\theta z| \geq |\Re(e^{\mathbf{i}\theta} z' P_\theta z)|$, we have:

$$\frac{|z'P_\theta z|}{\|z\|^2} \geq \frac{\sum r_j \max\{0, \cos(\theta + \phi_j)\}}{\sum r_j}$$

The numerator of the right member of the inequality above is a continuous function of θ maximized for some unknown value of θ, where $\theta/2$ is the phase of the free rotation. Although the maximum is difficult to locate, it cannot be smaller than the mean value of the function. This mean value is:

$$(1/2\pi) \int_{\theta=0}^{\theta=2\pi} \sum r_j \max\{0, \cos(\theta + \phi_j)\} d\theta$$

Permuting the two summation operators, we are left with a finite sum of integrals, each of them being equal to $2r_j$. The mean value of the function is $2\sum r_j/2\pi$, and thus we obtain Lemma 4:

Lemma 4. For any complex vector having at least half of its elements null, we have the following inequality: $[\max_{\{P\}} |z'Pz|/\|z\|^2] \geq 1/\pi$.

The condition "at least half of the elements are null" is asymptotically satisfied for the sets considered in the previous section. Now we remove this condition, and we set instead the centering condition $\mathbf{1}'z = 0$.

We know that an angle θ can be found such that:

$$\sum r_j \max\{0, \cos(\theta + \phi_j)\} \geq (1/\pi) \sum r_j$$

Let k be the number of elements z_j such that $\Re(z_j^2) < 0$. We define z_k as the k-dimensional vector such that $\Re(z_j^2) < 0$, and z_{n-k} as the $(n-k)$-dimensional vector such that $\Re(z_j^2) \geq 0$, such that $\mathbf{1}'z_k + \mathbf{1}'z_{n-k} = 0$. We set the arbitrary phase such that $\theta = 0$, without loss of generality:

$$\sum \max\{0, \Re(z_j^2)\} \geq (1/\pi) \sum |z_j|^2$$

Then we build the matrix $[nW_+]$, such that $[W_+]$ is a joint density matrix and $[nW_+]$ is a doubly stochastic matrix, as follows:

$$(n+k)[nW_+] = \mathbf{1} \cdot \mathbf{1}' + \left(\begin{array}{c|c} -\mathbf{1} \cdot \mathbf{1}' + n\mathbf{I} & 0 \\ \hline 0 & \mathbf{1} \cdot \mathbf{1}' \end{array}\right)$$

in which $\mathbf{I}$ is the identity matrix of size $n-k$ and the vectors $\mathbf{1}$ have the appropriate size (either k, or $n-k$, or n). Then, building $z' = [z'_{n-k}|z'_k]$, we have:

$$z'[nW_+]z = \frac{n}{n+k}(z'_{n-k}z_{n-k})$$

$$|z'[nW_+]z| \geq \frac{n}{n+k}|\Re(z'_{n-k}z_{n-k})|$$

$$|z'[nW_+]z| \geq \frac{n}{n+k}(\frac{1}{\pi}) \sum |z_j|^2$$

The permutation matrices are the extreme points of the closed bounded convex set of bistochastic matrices. Then, $\max_{\{P\}} |z'Pz| \geq |z'[nW_+]z|$ and:

$$\max_{\{P\}} |z'Pz| \geq \frac{n}{n+k}(\frac{1}{\pi})\|z\|^2$$

And since $k \leq n$:

$$\max_{\{P\}} |z'Pz|/\|z\|^2 \geq 1/2\pi$$

We obtain Lemma 5:

Lemma 5. The chiral index of any bivariate sample cannot be greater than $1 - 1/2\pi$.

A slight improvement is obtained when the condition $z'z = 0$ is added. We build the doubly stochastic matrix $[nW_-]$:

$$(n + (n-k))[nW_-] = \mathbf{1}\cdot\mathbf{1}' + \left(\begin{array}{c|c} \mathbf{1}\cdot\mathbf{1}' & 0 \\ \hline 0 & -\mathbf{1}\cdot\mathbf{1}' + n\mathbf{I} \end{array}\right)$$

Then:

$$z'[nW_-]z = \frac{n}{n+(n-k)}(z_k'z_k)$$

Since $z_k'z_k = -z_{n-k}'z_{n-k}$, we are led to the same inequalities as above, except that the factor $n/(n+k)$ is now replaced by $n/(n+(n-k))$:

$$\max_{\{P\}} |z'Pz| \geq \frac{n}{n+(n-k)}(\frac{1}{\pi})\|z\|^2$$

Depending which of k or $(n-k)$ is the smaller, the largest of the ratios $n/(n+k)$ and $n/(n+(n-k))$ cannot be smaller than $2/3$, and thus:

$$\max_{\{P\}} |z'Pz|/\|z\|^2 \geq 2/3\pi$$

The condition $z'z = 0$, i.e. $x'x = y'y$ and $x'y = 0$, means that the variance matrix of the centered set $[x|y]$ is proportional to the identity matrix. This condition is asymptotically satisfied by the sets described in the previous section.

6. Conclusion

From Lemmas 3 and 5, the upper bound of the chiral index of any bivariate sample is lying somewhere in the interval $[1 - 1/\pi, 1 - 1/2\pi]$. From the convergence theorem in section IV in [1], we deduce that this interval is still valid for any bivariate distribution with finite and non null inertia.

REFERENCES

[1] Petitjean, M., Chiral Mixtures, J. Math. Phys. 43 (2002) 4147-4157.

[2] Petitjean, M., Chirality and Symmetry Measures: A Transdisciplinary Review, Entropy 5 (2003) 271-312. *(http://www.mdpi.net/entropy)*

[3] Rachev, S.T., *Probability Metrics and the Stability of Stochastic Models*, Wiley, New-York, 1991.

[4] Petitjean, M., From Shape Similarity to Shape Complementarity: toward a Docking Theory, J. Math. Chem. 35 (2004) 147-158.

[5] Petitjean, M., About Second Kind Continuous Chirality Measures. 1.Planar Sets, J. Math. Chem. 22 (1997) 185-201.