Chapter 25
Distances Defined by Zonoids

The goals of this chapter are to:

- Introduce $\mathfrak{N}$-distances defined by zonoids,
- Explain the connections between $\mathfrak{N}$-distances and zonoids.

Notation introduced in this chapter:

Notation	Description
$h(K, u)$	Support function of a convex body
$K_1 \oplus K_2$	Minkowski sum of sets K_1 and K_2
S^{d-1}	Unit sphere in $\mathbb{R}^d$

25.1 Introduction

Suppose that $\mathfrak{X}$ is a metric space with the distance ρ. It is well known (Schoenberg 1938) that $\mathfrak{X}$ is isometric to a subspace of a Hilbert space if and only if ρ^2 is a negative definite kernel. The so-called $\mathfrak{N}$-distance (Klebanov 2005) is a variant of a construction of a distance on a space of measures on $\mathfrak{X}$ such that $\mathfrak{N}^2$ is a negative definite kernel. Such a construction is possible if and only if ρ^2 is a strongly negative definite kernel on $\mathfrak{X}$.

In this chapter, we show that the supporting function of any zonoid in $\mathbb{R}^d$ is a negative definite first-degree homogeneous function. The inverse is also true. If the support of a generating measure of a zonoid coincides with the unit sphere, then the supporting function is strongly negative definite, and therefore it generates a distance on the space of Borel probability measures on $\mathbb{R}^d$.

S.T. Rachev et al., *The Methods of Distances in the Theory of Probability and Statistics*, DOI 10.1007/978-1-4614-4869-3_25, © Springer Science+Business Media, LLC 2013

25.2 Main Notions and Definitions

Here we review some known definitions and facts from stochastic geometry.[1]

Let $\mathfrak{C}$ (resp. $\mathfrak{C}'$) be the system of all compact convex sets (resp. nonempty compact convex sets) in $\mathbb{R}^d$. A set $K \in \mathfrak{C}'$ is called a convex body if $K \in \mathfrak{C}'$; then for each $u \in S^{d-1}$ there is exactly one number $h(K,u)$ such that the hyperplane

$$\{x \in \mathbb{R}^d : \langle x,u\rangle - h(K,u) = 0\} \tag{25.2.1}$$

intersects K, and $\langle x,u\rangle - h(K,u) \le 0$ for each $x \in K$. This hyperplane is called the *support hyperplane*, and the function $h(K,u)$, $u \in S^{d-1}$ (where S^{d-1} is the unit sphere), is the *support function* (restricted to S^{d-1}) of K. Equivalently, one can define

$$h(K,u) = \sup\{\langle x,u\rangle,\ x \in K\},\quad u \in \mathbb{R}^d. \tag{25.2.2}$$

Its geometrical meaning is the signed distance of the support hyperplane from the coordinate origin.

An important property of $h(K,u)$ is its additivity:

$$h(K_1 \oplus K_2, u) = h(K_1,u) + h(K_2,u),$$

where $K_1 \oplus K_2 = \{a+b :\ a \in K_1,\ b \in K_2\}$ is the Minkowski sum of K_1 and K_2. For $K \in \mathfrak{C}'$ let $\check{K} = \{-k, k \in K\}$. We say that K is *centrally symmetric* if $K' = \check{K}'$ for some translate K', i.e., if K has a center of symmetry.

The Minkowski sum of finitely many centered line segments is called a *zonotope*. Consider a zonotope

$$\mathcal{Z} = \bigoplus_{i=1}^{k} a_i[v_i, -v_i], \tag{25.2.3}$$

where $a_i > 0$, $v_i \in \mathbb{S}^{d-1}$. Its support function is given by

$$h(\mathcal{Z},u) = h_{\mathcal{Z}}(u) = \sum_{i=1}^{k} a_i |\langle u, v_i\rangle|. \tag{25.2.4}$$

We use the notation $\mathcal{K}'$ for the space of all compact subsets of $\mathbb{R}^d$ with the Hausdorff metric

$$d_H(K_1,K_2) = \max\{\sup_{x\in K_1} \operatorname{dist}(x,K_2),\ \sup_{y\in K_2} \operatorname{dist}(y,K_1)\}, \tag{25.2.5}$$

where $\operatorname{dist}(x,K) = \inf_{z\in K} \|x-z\|$.

[1] See, for example, Ziegler (1995) and Beneš and Rataj (2004).

A set $\mathcal{Z} \in \mathfrak{C}'$ is called a *zonoid* if it is a limit in a d_H distance of a sequence of zonotopes.

It is known that a convex body $\mathcal{Z}$ is a zonoid if and only if its support function has a representation

$$h(\mathcal{Z}, u) = \int_{\mathbb{S}^{d-1}} |\langle u, v\rangle| \mathrm{d}\mu_{\mathcal{Z}}(v) \tag{25.2.6}$$

for an even measure $\mu_{\mathcal{Z}}$ on $\mathbb{S}^{d-1}$. The measure $\mu_{\mathcal{Z}}$ is called the *generating measure* of $\mathcal{Z}$. It is known that the generating measure is unique for each zonoid $\mathcal{Z}$.

25.3 $\mathfrak{N}$-Distances

Suppose that $(\mathfrak{X}, \mathfrak{A})$ is a measurable space and $\mathcal{L}$ is a strongly negative definite kernel on $\mathfrak{X}$. Denote by $\mathcal{B}_{\mathcal{L}}$ the set of all probabilities μ on $(\mathfrak{X}, \mathfrak{A})$ for which there exists the integral

$$\int_{\mathfrak{X}} \int_{\mathfrak{X}} \mathcal{L}(x, y) \mathrm{d}\mu(x) \mathrm{d}\mu(y) < \infty. \tag{25.3.1}$$

For $\mu, \nu \in \mathcal{B}_{\mathcal{L}}$ consider

$$\begin{aligned} \mathcal{N}(\mu, \nu) = 2 &\int_{\mathfrak{X}} \int_{\mathfrak{X}} \mathcal{L}(x, y) \mathrm{d}\mu(x) \mathrm{d}\nu(y) \\ &- \int_{\mathfrak{X}} \int_{\mathfrak{X}} \mathcal{L}(x, y) \mathrm{d}\mu(x) \mathrm{d}\mu(y) \\ &- \int_{\mathfrak{X}} \int_{\mathfrak{X}} \mathcal{L}(x, y) \mathrm{d}\nu(x) \mathrm{d}\nu(y). \end{aligned} \tag{25.3.2}$$

It is known (Klebanov 2005) that

$$\mathfrak{N}(\mu, \nu) = \left(\mathcal{N}(\mu, \nu)\right)^{1/2}$$

is a distance on $\mathcal{B}_{\mathcal{L}}$.

Described below are some examples of negative definite kernels.

Example 25.3.1. Let $\mathfrak{X} = \mathbb{R}^1$. For $r \in [0, 2]$ define

$$\mathcal{L}_r(x, y) = |x - y|^r.$$

The function $\mathcal{L}_r$ is a negative definite kernel. For $r \in (0, 2)$, $\mathcal{L}_r$ is a strongly negative definite kernel.

For the proof of the statement in this example and the statement in the next example (Example 25.3.2), see Klebanov (2005).

Example 25.3.2. Let $\mathcal{L}(x, y) = f(x - y)$, where $f(t)$ is a continuous function on $\mathbb{R}^d$, $f(0) = 0$, $f(-t) = f(t)$. $\mathcal{L}$ is a negative definite kernel if and only of

$$f(t) = \int_{\mathbb{R}^d} \big(1 - \cos\langle t, u\rangle\big) \frac{1 + \|u\|^2}{\|u\|^2} \mathrm{d}\Theta(u), \tag{25.3.3}$$

where Θ is a finite measure on $\mathbb{R}^d$. Representation (25.3.3) is unique. Kernel $\mathcal{L}$ is strongly negative definite if the support of the measure Θ coincides with the whole space $\mathbb{R}^d$.

We will give an alternative proof for the fact that $|x - y|$ is a negative definite kernel. For the case $\mathfrak{X} = \mathbb{R}^1$ define

$$\mathcal{L}(x, y) = 2 \max(x, y) - x - y = |x - y|. \tag{25.3.4}$$

Then $\mathcal{L}$ is a negative definite kernel.

Proof. It is sufficient to show that $\max(x, y)$ is a negative definite kernel. For arbitrary $a \in \mathbb{R}^1$ consider

$$u_a(x) = \begin{cases} 1, & x < a, \\ 0, & x \geq a. \end{cases} \tag{25.3.5}$$

It is clear that

$$u_a(\max(x, y)) = u_a(x)u_a(y).$$

Let $F(a)$ be a nondecreasing bounded function on $\mathbb{R}^1$. Define

$$\mathcal{K}(x, y) = \int_{-\infty}^{\infty} u_a(\max(x, y))\mathrm{d}F(a).$$

For any integer $n > 1$ and arbitrary $c_1, \ldots, c_n$ under condition $\sum_{j=1}^n c_j = 0$ we have

$$\begin{aligned} \sum_{i=1}^n \sum_{j=1}^n \mathcal{K}(x_i, x_j)c_i c_j &= \int_{-\infty}^{\infty} \sum_{i=1}^n \sum_{j=1}^n u_a(x_i)u_a(x_j)c_i c_j \mathrm{d}F(a) \\ &= \int_{-\infty}^{\infty} \left(\sum_{i=1}^n u_a(x_i)c_i \right)^2 \mathrm{d}F(a) \geq 0. \end{aligned}$$

But

$$\mathcal{K}(x,y) = \int_{-\infty}^{\infty} u_a(\max(x,y))\mathrm{d}F(a)$$
$$= F(+\infty) - F(\max(x,y)).$$

Let us fix arbitrary $A > 0$ and apply the previous equality to the function

$$F(a) = F_A(a) = \begin{cases} A & \text{for } a > A, \\ a & \text{for } -A \le a \le A, \\ -A & \text{for } a < -A. \end{cases} \quad (25.3.6)$$

In this case, $\mathcal{K}(x,y) = A - \max(x,y)$ for $x, y \in [-A, A]$, and, as $A \to \infty$, we obtain that $\max(x,y)$ is a negative definite kernel. □

Directly from the definition of a negative definite kernel and Example 25.3.1 we obtain the next example.

Example 25.3.3. Let $x, y \in \mathbb{R}^d$, and $f : \mathbb{R}^d \to \mathbb{R}^1$. Define

$$\mathcal{L}(x,y) = |f(x) - f(y)|.$$

Then $\mathcal{L}$ is a negative definite kernel.

Of course, the mixture of negative definite kernels is again a negative definite kernel.

Example 25.3.4. Let us choose and fix a vector $\theta \in \mathbb{S}^{d-1}$ and consider the kernel

$$\mathcal{L}_\theta(x,y) = |\langle x,\theta\rangle - \langle y,\theta\rangle|.$$

From previous considerations it is clear that $\mathcal{L}_\theta$ is a negative definite kernel on $\mathbb{R}^d$, and for the σ-finite measure Ξ

$$\mathcal{L}_\Xi(x,y) = \int_{\mathbb{S}^{d-1}} \mathcal{L}_\theta(x,y)\mathrm{d}\Xi(\theta) \quad (25.3.7)$$

is, again, a negative definite kernel.

Consider expression (25.3.2) constructed on the basis of (25.3.7). Let us rewrite (25.3.2) in a different form. Suppose that X and Y are two random vectors in $\mathbb{R}^d$ with distributions μ and ν, respectively. We write $\mathcal{N}(X,Y)$ instead of $\mathcal{N}(\mu,\nu)$, so that

$$\mathcal{N}(X,Y) = 2E\mathcal{L}_\Xi(X,Y) - E\mathcal{L}_\Xi(X,X') - E\mathcal{L}_\Xi(Y,Y'),$$

where $X' \stackrel{d}{=} X$ and $Y' \stackrel{d}{=} Y$ are independent copies of X and Y, respectively. Note that we use the sign $\stackrel{d}{=}$ for the equality in a distribution. We have

$$\mathcal{N}(X,Y) = E\int_{\mathbb{S}^{d-1}} [4\max(\langle X,\theta\rangle, \langle Y,\theta\rangle) \\ -2\max(\langle X,\theta\rangle, \langle X',\theta\rangle) - 2\max(\langle Y,\theta\rangle, \langle Y',\theta\rangle)]\mathrm{d}\Xi(\theta).$$

Denote $X_\theta = \langle X,\theta\rangle$, $Y_\theta = \langle Y,\theta\rangle$. Then

$$\mathcal{N}(X,Y) = 2\int_{\mathbb{S}^{d-1}} \lim_{A\to\infty} E\int_{-A}^{A} \big(u_a(X_\theta)u_a(X'_\theta) \\ +u_a(Y_\theta)u_a(Y'_\theta) - 2u_a(X_\theta)u_a(Y_\theta)\big)\mathrm{d}F_A(a)\mathrm{d}\Xi(\theta).$$

But $Eu_a(X_\theta) = \Pr\{X_\theta < a\}$, and therefore

$$\mathcal{N}(X,Y) = 2\lim_{A\to\infty}\int_{\mathbb{S}^{d-1}} \mathrm{d}\Xi(\theta)\int_{-A}^{A} \Big(\Pr\{X_\theta < a\}\Pr\{X'_\theta < a\} \\ +\Pr\{Y_\theta < a\}\Pr\{Y'_\theta < a\} - 2\Pr\{X_\theta < a\}\Pr\{Y_\theta < a\}\Big)\mathrm{d}F_A(a) \\ = 2\int_{\mathbb{S}^{d-1}} \mathrm{d}\Xi(\theta)\int_{-\infty}^{\infty}\Big(F_\theta(a) - G_\theta(a)\Big)^2 \mathrm{d}a,$$

where $F_\theta(a) = \Pr\{X_\theta < a\}$, $G_\theta(a) = \Pr\{Y_\theta < a\}$. So finally we have

$$\mathcal{N}(X,Y) = 2\int_{\mathbb{S}^{d-1}} \mathrm{d}\Xi(\theta)\int_{-\infty}^{\infty}\Big(F_\theta(a) - G_\theta(a)\Big)^2 \mathrm{d}a. \tag{25.3.8}$$

If the support of Ξ coincides with $\mathbb{S}^{d-1}$, then $\mathfrak{N}(X,Y) = \Big(\mathcal{N}(X,Y)\Big)^{1/2}$ is a distance between the distributions of X and Y.

Let us return to the kernel

$$\mathcal{L}_\theta(x,y) = 2\max(\langle x,\theta\rangle, \langle y,\theta\rangle) - \langle x,\theta\rangle - \langle y,\theta\rangle.$$

Choose arbitrary $\theta_o \in \mathbb{S}^{d-1}$, and consider the measure

$$\Xi_o = \frac{1}{2}\big(\delta_{\theta_o} + \delta_{-\theta_o}\big),$$

where δ_{θ_o} is the measure concentrated at point θ_o. Then

$$\begin{aligned}\mathcal{L}_{\Xi_{\theta_o}}(x,y) &= \int_{\mathbb{S}^{d-1}} \mathcal{L}_\theta(x,y)\mathrm{d}\Xi_o(\theta)\\ &= \max(\langle x,\theta_o\rangle,\langle y,\theta_o\rangle) + \max(-\langle x,\theta_o\rangle,-\langle y,\theta_o\rangle)\\ &= |\langle x-y,\theta\rangle|.\end{aligned}$$

Now, if we have an arbitrary even measure Ξ_s on sphere $\mathbb{S}^{d-1}$, then

$$\begin{aligned}\mathcal{L}_{\Xi_s}(x,y) &= \int_{\mathbb{S}^{d-1}} \mathcal{L}_\theta(x,y)\mathrm{d}\Xi_s(\theta)\\ &= \int_{\mathbb{S}^{d-1}} |\langle x-y,\theta\rangle|\mathrm{d}\Xi_s(\theta)\end{aligned}$$

is a negative definite kernel. Let us note that the function

$$h(z) = \int_{\mathbb{S}^{d-1}} |\langle z,\theta\rangle|\mathrm{d}\Xi_s(\theta), \quad z\in\mathbb{R}^d \tag{25.3.9}$$

is the support function of a zonoid with generating measure Ξ_s.

Summarizing all the preceding relations we may formulate the following result.

Theorem 25.3.1. Each zonoid $\mathcal{Z}$ generates a negative definite kernel on $\mathbb{R}^d$

$$\mathcal{L}_{\mathcal{Z}}(x,y) = h_{\mathcal{Z}}(x-y) = \int_{\mathbb{S}^{d-1}} |\langle x-y,\theta\rangle|\mathrm{d}\mu_{\mathcal{Z}}(\theta). \tag{25.3.10}$$

This kernel is strongly negative definite if the support of $\mu_{\mathcal{Z}}$ coincides with the whole sphere $\mathbb{S}^{d-1}$, and

$$\begin{aligned}\mathcal{N}(\mu,\nu) = 2\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \mathcal{L}_{\mathcal{Z}}(x,y)\mathrm{d}\mu(x)\mathrm{d}\nu(y) - \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \mathcal{L}_{\mathcal{Z}}(x,y)\mathrm{d}\mu(x)\mathrm{d}\mu(y)\\ - \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \mathcal{L}_{\mathcal{Z}}(x,y)\mathrm{d}\nu(x)\mathrm{d}\nu(y)\end{aligned}$$

is the square of a distance between measures $\mu,\nu \in \mathcal{B}_{\mathcal{L}}$. This distance has the following representation:

$$\mathfrak{N}(\mu,\nu) = \left(\int_{\mathbb{S}^{d-1}} \mathrm{d}\mu_{\mathcal{Z}}(\theta)\int_{-\infty}^{\infty} \big(F_\theta(a) - G_\theta(a)\big)^2 \mathrm{d}a\right)^{1/2}, \tag{25.3.11}$$

where

$$\begin{aligned}&\mu(\mathcal{A}) = \Pr\{X\in\mathcal{A}\},\ \nu(\mathcal{A}) = \Pr\{Y\in\mathcal{A}\},\\ &F_\theta(a) = \Pr\{\langle X,\theta\rangle < a\},\ G_\theta(a) = \Pr\{\langle Y,\theta\rangle < a\}.\end{aligned} \tag{25.3.12}$$

According to Example 25.3.2, the function $h_{\mathcal{Z}}(u)$ from (25.3.10) may be represented in the form (25.3.3). Let us investigate the connection between $\mu_{\mathcal{Z}}$ in (25.3.10) and Θ in (25.3.3). To do so, we will use the following identity:

$$|z| = \frac{2}{\pi}\int_0^\infty \big(1-\cos(zt)\big)\frac{\mathrm{d}t}{t^2}. \tag{25.3.13}$$

We have

$$\begin{aligned} h_{\mathcal{Z}}(u) &= \frac{2}{\pi}\int_{\mathbb{S}^{d-1}}\int_0^\infty \big(1-\cos\langle u,\theta\rangle\big)\frac{\mathrm{d}t}{t^2}\mathrm{d}\mu_{\mathcal{Z}}(\theta) \\ &= \frac{2}{\pi}\int_{\mathbb{R}^d}\big(1-\cos\langle u,v\rangle\big)\frac{1+\|v\|^2}{\|v\|^2}\mathrm{d}\Theta(v). \end{aligned}$$

So

$$\begin{aligned} \mathrm{d}\Theta(v) &= \frac{2}{\pi}\frac{1}{1+t^2}\mathrm{d}t\,\mathrm{d}\mu(\theta), \\ v &= t\cdot\theta,\quad \theta\in\mathbb{S}^{d-1},\ t\geq 0. \end{aligned} \tag{25.3.14}$$

If $h_{\mathcal{Z}}(u)$ is a support function of a zonoid $\mathcal{Z}$, then clearly

$$h_{\mathcal{Z}}(\tau\cdot u) = \tau h_{\mathcal{Z}}(u)$$

for all $\tau > 0$ and $u \in \mathbb{R}^d$, and, as was shown previously, $h_{\mathcal{Z}}(x-y)$ is a negative definite kernel. The inverse is also true.

Theorem 25.3.2. Suppose that f is a continuous function on $\mathbb{R}^d$ such that $f(0) = 0$, $f(-u) = f(u)$. Then the following facts are equivalent:

Fact 1. $f(\tau\cdot u) = \tau f(u)$ and $f(x-y)$ is a negative definite kernel.
Fact 2. f is a support function of a zonoid.

Proof. Previously we saw that Fact 2 implies Fact 1, and we must prove only that Fact 1 implies Fact 2. According to Example 25.3.2,

$$f(u) = \int_{\mathbb{R}^d}\big(1-\cos\langle u,v\rangle\big)\mathrm{d}\Theta_1(v), \tag{25.3.15}$$

where

$$\mathrm{d}\Theta_1(v) = \frac{1+\|v\|^2}{\|v\|^2}\mathrm{d}\Theta(v),$$

and Θ is the measure from (25.3.3).

We have

$$f(\tau\cdot u) = \tau f(u) \tag{25.3.16}$$

for any $\tau > 0$, $u \in \mathbb{R}^d$. Substituting (25.3.15) into (25.3.16) and using the uniqueness of the measure Θ in (25.3.3) we obtain

$$\int_{\mathbb{R}^d} \big(1 - \cos\langle \tau \cdot u, v\rangle\big) \mathrm{d}\Theta_1(v) = \tau \int_{\mathbb{R}^d} \big(1 - \cos\langle u, v\rangle\big) \mathrm{d}\Theta_1(v),$$

$$\big(1 - \cos\langle u, v\rangle\big) \mathrm{d}\Theta_1(v/\tau) = \tau \int_{\mathbb{R}^d} \big(1 - \cos\langle u, v\rangle\big) \mathrm{d}\Theta_1(v)$$

and

$$\Theta_1(v/\tau) = \tau\Theta_1(v).$$

We write here $v = r \cdot w$ for $r > 0$ and $w \in \mathbb{S}^{d-1}$. We have

$$\Theta_1(r\tau \cdot w) = \tau\Theta_1(r \cdot w)$$

and, finally, for $\tau = r$,

$$\Theta_1(r \cdot w) = \frac{1}{r}\Theta_1(w). \tag{25.3.17}$$

It is clear that representation (25.3.15) for Θ_1 of the form (25.3.17) coincides with (25.3.14).[2] □

Note that the $\mathfrak{N}$-distance can be bounded by the Hausdorf distance. Let $\mathcal{Z}_\mu$ and $\mathcal{Z}_\nu$ be two zonoids with generating measures μ and ν, respectively. The following inequality holds for their supporting functions $h(\mathcal{Z}_\mu, u)$ and $h(\mathcal{Z}_\nu, u)$:

$$|h(\mathcal{Z}_\mu, u) - h(\mathcal{Z}_\nu, u)| \le d_H(\mathcal{Z}_\mu, \mathcal{Z}_\nu).$$

Obviously, from this inequality it follows that

$$\mathcal{N}(\mu, \nu) \le 2d_H(\mathcal{Z}_\mu, \mathcal{Z}_\nu),$$

and therefore

$$\mathfrak{N}(\mu, \nu) \le (2d_H(\mathcal{Z}_\mu, \mathcal{Z}_\nu))^{1/2}. \tag{25.3.18}$$

Note that each $\mathfrak{N}$-distance generated by a zonoid is an ideal distance of degree $1/2$.

References

Beneš V, Rataj J (2004) Stochastic geometry: selected topics. Kluwer, Boston
Burger M (2000) Zonoids and conditionally positive definite functions. Portugaliae Mathematica 57:443–458
Klebanov LB (2005) $\mathfrak{N}$-distances and their applications. Karolinum Press, Prague
Schoenberg IJ (1938) Metric spaces and positive definite functions. Trans Am Math Soc 44:552–563
Ziegler GM (1995) Lectures on polytopes. Springer, New York

[2] An alternative proof of Theorem 25.3.2 is provided in Burger (2000).