Contents

1 Statistical Methods IV: Depth Statistics

Course Information

- **Instructor:** Shyamal K De
- **Department:** ASU
- **Email:** skd.isical@gmail.com
- **Marking Scheme:** Class Test 10%, Project 15%, Midterm 25%, End Semester 50%

Author Information

- **Name:** Mohammad Shaan
- **Academic Year:** B.Stat II Year
- **Email:** mdsworld2006@.com

1.1 Foundational axioms of depth statistics

This section introduces the **foundational axioms of depth statistics**. A *depth function* assigns to a point x a numerical value measuring how *central* it is relative to a distribution or data cloud.

The framework here is conceptual rather than computational. The goal is to formalize what counts as a legitimate notion of "centrality," independent of any particular construction.

Concretely, the authors are doing three things:

- giving a **general abstract definition** of depth;
- stating **axioms (postulates)** such a function must satisfy;
- explaining how these axioms encode geometric, probabilistic, and robustness properties.

1.1.1 Mathematical setting

Space and probability structure

- **State space:** E is a **Banach space**, i.e. a complete normed vector space.
  - $\mathbb{R}^d$ with the Euclidean norm;

  - function spaces such as L^p.
- **Sigma-algebra:** $\mathcal{B}$ is the **Borel σ-algebra** on E, generated by all open sets.
- **Distributions:** $\mathcal{P}$ is a class of probability measures on $(E, \mathcal{B})$.
  - often all probability measures;
  - sometimes restricted (e.g. finite moments).
- **Random element:** X is an E-valued random variable with distribution $P \in \mathcal{P}$.

1.1.2 Definition: depth function

Definition 1.1. A **depth function** is a mapping

$$D : E \times \mathcal{P} \to [0, 1], \qquad (x, P) \mapsto D(x \mid P).$$

Interpretation.

- $D(x \mid P)$ measures how *central* the point x is relative to P;
- larger values indicate greater centrality;
- the range $[0, 1]$ is a normalization convention.

1.1.3 Postulates D1–D5

These are **axioms**, not derived properties. A function that violates them is not considered a valid depth within this framework.

D1. Translation invariance

Statement.

$$D(x + b \mid X + b) = D(x \mid X), \quad \forall b \in E.$$

Meaning. Shifting both the point and the distribution by the same vector leaves depth unchanged.

Rationale. Depth depends on relative position, not absolute coordinates.

D2. Linear invariance

Statement.

$$D(Ax \mid AX) = D(x \mid X),$$

for every **bijective linear transformation** $A : E \to E$.

Meaning. Depth is invariant under changes of units, rotations, reflections, and shearings.

Why bijectivity matters. Non-invertible maps destroy information and cannot preserve centrality.

D3. Null at infinity

Statement.

$$\lim_{|x|\to\infty} D(x \mid X) = 0.$$

Meaning. Points arbitrarily far from the data cloud have negligible depth.

Role. Excludes pathological definitions assigning large depth far away from the data.

D4. Monotonicity on rays

Statement. If x^* is a point of maximal depth, then for any unit vector r,

$$D(x^* + \alpha r \mid X) \quad \text{decreases as } \alpha > 0 \text{ increases.}$$

Meaning. Moving away from a deepest point in any direction reduces centrality.

Role. Prevents oscillatory or non-radial depth behavior.

D5. Upper semicontinuity

Statement. For every $\alpha \in (0, 1]$,

$$\{x \in E : D(x \mid X) \geq \alpha\} \quad \text{is closed.}$$

Consequences.

- ensures existence of maximizers;
- guarantees stability under small perturbations.

Upper semicontinuity rules out sudden upward jumps of depth.

1.1.4 Consequences of D1–D4

- D1 and D2 imply **affine invariance**;
- D3 and D4 imply that upper level sets are **bounded and star-shaped** around deepest points.

Star-shaped sets

A set S is star-shaped about x^* if

$$x \in S \Rightarrow [x^*, x] \subset S.$$

This guarantees that central regions are solid, connected regions rather than disconnected shells.

1.1.5 Proposition 2.1

Proposition 1.2. *If X is centrally symmetric about x^*, then any depth function satisfying D1–D5 is maximized at x^*.*

Central symmetry

$$X - x^* \overset{d}{=} x^* - X.$$

This proposition ensures that depth correctly identifies the center under symmetry.

1.1.6 Strengthened and weakened axioms

D4con (quasiconcavity)

Upper level sets of depth are **convex**.

Depths satisfying this condition are called **convex depths**. Convexity implies star-shapedness, so D4con is stronger than D4.

D2iso (isometric invariance)

$$D(Ax \mid AX) = D(x \mid X)$$

only for **isometries**, i.e. distance-preserving maps.

This weakens affine invariance to preserve scale and variance structure.

D2sca (scale invariance)

$$D(\lambda x \mid \lambda X) = D(x \mid X), \quad \forall \lambda > 0.$$

Depth should not depend on the units of measurement.

1.2 Central Regions and Outliers

Restated goal of the section

This section explains how a depth function induces:

- **central regions** (multivariate analogues of quantiles);
- a notion of **outlyingness**;
- an ordering of points from "most central" to "most extreme".

This forms the bridge between abstract depth axioms and practical data analysis.

1.2.1 Step 1: Level sets of a depth function

Given objects

- $P \in \mathcal{P}$: a probability distribution on E;

- $D(\cdot \mid P)$: a depth function satisfying D1–D5;
- $\alpha \in [0, 1]$.

Definition: α-central region

$$D_\alpha(P) \;:=\; \{x \in E : D(x \mid P) \geq \alpha\}.$$

These are also called **depth-trimmed regions** or **central regions**.

Formal properties

From axioms D3–D5 and possibly D4con:

- closed (D5);
- bounded (D3);
- star-shaped (D4);
- convex if D4con holds.

Nested structure

If $0 \leq \alpha_1 < \alpha_2 \leq 1$, then

$$D_{\alpha_2}(P) \subseteq D_{\alpha_1}(P).$$

Higher depth corresponds to a stricter inclusion criterion.

1.2.2 Step 2: The deepest point(s)

Define

$$\alpha_{\max} := \sup_{x \in E} D(x \mid P).$$

Then $D_{\alpha_{\max}}(P)$ is the set of **deepest points**.

- If unique, it is a multivariate median;
- otherwise, it forms a deepest region.

1.2.3 Step 3: Affine equivariance of central regions

From D1 and D2, if A is affine and bijective, then

$$D_\alpha(AP) = AD_\alpha(P).$$

Central regions transform exactly like the data, preserving geometry such as location, spread, and shape.

1.2.4 Step 4: From regions to depth

Assumption

We are given a nested family of sets $\{C_\alpha(P)\}_{\alpha \in [0,1]}$, each of which is convex, bounded, closed, and nested in α.

Definition (Equation 2.1)

$$D(z \mid P) \;:=\; \sup\{\alpha : z \in C_\alpha(P)\}. \tag{2.1}$$

Formal justification

- nestedness ensures the supremum is well defined;
- closedness implies upper semicontinuity;
- convexity yields quasiconcavity.

This construction automatically satisfies D5 and D4con.

1.2.5 Step 5: Depth ordering and outlyingness

Center–outward ordering

$$D(x \mid X) > D(y \mid X) \;\Rightarrow\; x \text{ is more central than } y.$$

Depth induces an ordering from the center to the periphery.

1.2.6 Step 6: Outlyingness function

Definition

$$\mathrm{Out}(z \mid X) \;:=\; \frac{1}{D(z \mid X)} - 1.$$

Properties

- if z is deepest: $D = 1 \Rightarrow \mathrm{Out} = 0$;
- if $D \downarrow 0$: $\mathrm{Out} \uparrow \infty$.

The transformation converts “large depth” into “small outlyingness”.

$$D(z \mid X) = \frac{1}{1 + \mathrm{Out}(z \mid X)}.$$

1.2.7 Step 7: Outliers via central regions

If

$$z \notin D_\alpha(X),$$

then

$$\mathrm{Out}(z \mid X) > \frac{1}{\alpha} - 1.$$

This yields a distribution-free notion of multivariate outliers.

1.3 Depth Lifts, Stochastic Orderings, and Metrics

Goal of the section

Depth induces:

- a partial ordering on distributions;
- a notion of dispersion;
- a distance (semi-metric) between distributions.

1.3.1 Step 1: Normalization

Assume

$$\alpha_{\max} = 1.$$

This is a harmless rescaling of depth.

1.3.2 Step 2: Definition of the depth lift

$$\widehat{D}(P) \;:=\; \{(\alpha, x) \in [0,1] \times E : x \in D_\alpha(P)\}. \tag{2.2}$$

This stacks all central regions along the α-axis into a single geometric object.

1.3.3 Step 3: Dispersion ordering of distributions

$$P \preceq_D Q \quad \text{if} \quad \widehat{D}(P) \subseteq \widehat{D}(Q). \tag{2.3}$$

Equivalently,

$$D_\alpha(P) \subseteq D_\alpha(Q) \quad \text{for all } \alpha.$$

This means that Q is more dispersed than P.

1.3.4 Step 4: Partial order vs preorder

- if the family $\{D_\alpha(P)\}$ uniquely determines P, then $\preceq_D$ is a partial order;
- otherwise, it is only a preorder.

1.3.5 Step 5: Depth-induced distance

$$\delta_D(P, Q) \;:=\; \delta_H(\widehat{D}(P), \widehat{D}(Q)), \tag{2.4}$$

where δ_H is the Hausdorff distance.

Hausdorff distance

$$\delta_H(C_1, C_2) = \inf\{\varepsilon > 0 : C_1 \subseteq C_2^\varepsilon \text{ and } C_2 \subseteq C_1^\varepsilon\}.$$

Two distributions are close if their entire depth geometry is close.

1.4 Multivariate Depth Functions

What this section is really about

Up to this point, depth has been defined **axiomatically**. Section 2.3 addresses the natural next question:

> *Which concrete depth functions actually exist, and how do they relate to the geometry of distributions?*

The central insight is that depth functions are not universal. Their behavior is tightly coupled to distributional shape, especially ellipticity, and to the invariance structure they satisfy.

1.4.1 Step 1: Why ellipticity matters

Definition

A random vector $X \in \mathbb{R}^d$ has a **spherical distribution** if

$$AX \overset{d}{=} X \quad \text{for every orthogonal matrix } A.$$

It has an **elliptical distribution** if

$$X = \mu + BY,$$

where:

- Y is spherical;
- B is a nonsingular $d \times d$ matrix;
- $\mu \in \mathbb{R}^d$.

This class includes multivariate normal, multivariate t, and many heavy-tailed models.

Geometric meaning

- spherical distributions have spherical level sets;
- elliptical distributions have ellipsoidal level sets.

This geometry determines what any affine invariant depth can detect.

1.4.2 Step 2: Depth under ellipticity

Assume:

- X is elliptically distributed;
- $D(\cdot \mid P)$ satisfies affine invariance (D2).

Key consequence

All depth contours $D_\alpha(P)$ are ellipsoids with the same shape matrix as the distribution.

Equivalently, depth reduces to a radial function:

$$D(x \mid P) = f\Big((x-\mu)'\Sigma^{-1}(x-\mu)\Big),$$

for some decreasing function f.

Why this is forced

- elliptical distributions are affine images of spherical ones;
- affine invariance forces depth contours to transform accordingly;
- spherical symmetry forces contours to be spheres;
- affine images of spheres are ellipsoids.

No additional assumptions are required.

Crucial implication

For elliptically distributed data, **all affine invariant depth functions induce the same ordering**.

They may differ numerically, but they rank points identically. In particular, Tukey depth, simplicial depth, Oja depth, and projection depth are equivalent as orderings under ellipticity.

1.4.3 Step 3: Consequences for inference

If P is elliptically symmetric and unimodal:

- all affine invariant depths identify the same center μ;
- all central regions have the same shape;
- all depths reduce to monotone functions of Mahalanobis distance.

This explains why classical Gaussian methods perform well under ellipticity and why depth methods differ meaningfully only outside this regime.

1.4.4 Step 4: Beyond ellipticity

For non-elliptical distributions:

- different depths produce different central regions;
- they respond differently to skewness and multimodality;
- they may or may not uniquely determine the distribution.

Depth functions then differ in robustness, sensitivity to asymmetry, geometry, and computational feasibility.

1.4.5 Step 5: Three main construction principles

The text identifies three fundamental ways to construct multivariate depths.

(I) Distance-based depths

Core idea: outlyingness is distance from a suitable center.

$$D(x) = g(\mathrm{dist}(x, \mathrm{center})) .$$

Centers may be means, medians, or deepest points; distances may be Euclidean, Mahalanobis, or robust.

These depths are intuitive and computationally simple, but often center-dependent and sensitive to shape misspecification.

(II) Weighted-mean depths

Core idea: central points balance the data.

$$D(x) = \left\| \sum_i w_i(x)(x_i - x) \right\|^{-1}.$$

This class includes spatial depth and L_1-type constructions. They are smooth and robust but may lack strong geometric guarantees or affine invariance.

(III) Simplicial and halfspace depths

Core idea: central points participate in many balanced geometric configurations.

- halfspace (Tukey) depth;
- simplicial depth;
- Oja depth.

These depths have strong geometric meaning and robustness, but are computationally expensive and have more complex asymptotics.

1.4.6 Step 6: No universally optimal depth

No depth function satisfies all desirable properties simultaneously. Trade-offs are unavoidable:

- robustness versus computability;
- smoothness versus exact geometry;
- affine invariance versus sensitivity to shape.

Depth theory is therefore a theory of principled choices rather than optimization.

1.4.7 Summary of common multivariate depth functions

1.5 L_2-Depth and the Depth Axioms

1.5.1 L_2-Depth and Axiom D1 (Translation Invariance)

Axiom D1: Formal statement

Let X be an $\mathbb{R}^d$-valued random vector with distribution P. Let $D(\cdot \mid X)$ be a statistical depth function.

Depth name	**Formal definition** $D(z \mid X)$	**Underlying intuition**	**Visualization principle**	**Associated median / deepest points**	**Claimed properties (no proof)**
L_2**-depth**	$\big(1 + \mathbb{E}\lvert z - X\rvert\big)^{-1}$	Mean outlyingness measured by average Euclidean distance	Inverse-distance contours; closer to data cloud implies higher depth	**Spatial median**: minimizers of $\mathbb{E}\lvert z - X\rvert$	D1, D3–D5; convex compact regions; not affine invariant; poor dispersion ordering
Affine-invariant L_2**-depth**	$\big(1 + \mathbb{E}\lvert z - X\rvert_{S_X}\big)^{-1}$	Distance scaled by affine-equivariant scatter	Ellipsoidal distortion correcting anisotropy	Affine spatial median	D1–D5; affine invariant; depends on choice of S_X
Mahalanobis depth	$\big(1 + (z - c_X)^\top S_X^{-1}(z - c_X)\big)^{-1}$	Squared standardized distance from center	Ellipsoids centered at c_X	Center c_X	D1, D2, D3–D5; moment-based; insensitive beyond first two moments
Moment Mahalanobis depth	$\big(1 + (z - \mathbb{E}X)^\top \Sigma_X^{-1}(z - \mathbb{E}X)\big)^{-1}$	Classical quadratic distance	Covariance ellipsoids	Mean vector $\mathbb{E}X$	Same as Mahalanobis; non-robust
Projection depth	$\Big(1 + \sup_{\lvert p\rvert=1} \frac{\langle p,z\rangle - \mathrm{med}(\langle p,X\rangle)}{\mathrm{Dmed}(\langle p,X\rangle)}\Big)^{-1}$	Worst standardized univariate outlyingness	Intersection of all directional slabs	Projection median (possibly non-unique)	D1–D5; D4con; high robustness
Oja depth	$\Big(1 + \frac{\mathbb{E}\,\mathrm{Vol}_d\big(\mathrm{co}(z,X_1,\ldots,X_d)\big)}{\sqrt{\det \Sigma_X}}\Big)^{-1}$	Average simplex volume with vertex at z	Volume grows rapidly away from center	**Oja median** (may be non-unique)	D1–D5; continuous; distribution-determining (compact support)
Location / Tukey depth	$\inf\{P(H) : H \text{ closed halfspace}, z \in H\}$	Minimum mass in any direction containing z	Nested halfspace intersections	Tukey median(s)	D1, D2, D3–D5; maximum depth $\leq 1/2$; strong robustness
Simplicial depth	$P\big(z \in \mathrm{co}(X_1, \ldots, X_{d+1})\big)$	Likelihood that z lies inside random simplices	Coverage frequency by simplices	Simplicial median(s)	D1, D2, D3–D5 (under continuity); sensitive to discreteness

Axiom D1 (Translation Invariance). For every $z \in \mathbb{R}^d$ and every translation vector $b \in \mathbb{R}^d$,

$$D(z \mid X) = D(z + b \mid X + b),$$

where $X + b := X + b$ almost surely.

Meaning of Axiom D1

Translation invariance requires that depth depends only on *relative position*, not on absolute location. Shifting both the data cloud and the point under inspection by the same vector leaves their geometric relationship unchanged, and depth must remain unchanged.

Why Axiom D1 is desirable

Without D1:
- the choice of origin would matter;
- identical data configurations at different locations would yield different depths;
- the notion of a center would be ill-defined.

Verification for the L_2-depth

Step 1: Definition. The L_2-depth of $z \in \mathbb{R}^d$ with respect to $X \sim P$ is

$$D^{L_2}(z \mid X) = \big(1 + \mathbb{E}|z - X|\big)^{-1}.$$

Step 2: Translate data and point. Fix $b \in \mathbb{R}^d$ and define $z' = z + b$, $X' = X + b$.

Step 3: Expand the translated depth.

$$D^{L_2}(z' \mid X') = \big(1 + \mathbb{E}|z' - X'|\big)^{-1}.$$

Step 4: Translation invariance of subtraction. For all $a, b, c \in \mathbb{R}^d$, $(a + c) - (b + c) = a - b$, hence

$$|(z + b) - (X + b)| = |z - X|.$$

Step 5: Expectations.

$$\mathbb{E}|z' - X'| = \mathbb{E}|z - X|.$$

Step 6: Conclusion.

$$D^{L_2}(z' \mid X') = D^{L_2}(z \mid X).$$

Geometric intuition. L_2-depth measures average distance from the data cloud. Shifting everything by the same vector preserves all distances; the definition depends only on differences $z - X$.

1.5.2 L_2-Depth and Axiom D2 (Linear Invariance)

Axiom D2: Formal statement

Let X be an $\mathbb{R}^d$-valued random vector. **Axiom D2 (Linear Invariance)** requires that for every $z \in \mathbb{R}^d$ and every invertible matrix $A \in \mathbb{R}^{d\times d}$,

$$D(z \mid X) = D(Az \mid AX).$$

Meaning and desirability

Linear invariance demands that depth be unaffected by rotations, reflections, scalings, or shearings. Together with D1 it yields affine invariance, making depth a geometric—not coordinate—property.

Verification for the L_2-depth

Definition.

$$D^{L_2}(z \mid X) = \big(1 + \mathbb{E}|z - X|\big)^{-1}.$$

Apply a linear map. Let A be invertible, $z' = Az$, $X' = AX$:

$$D^{L_2}(z' \mid X') = \big(1 + \mathbb{E}|A(z - X)|\big)^{-1}.$$

Norm distortion. The Euclidean norm satisfies $|Av| = |v|$ iff A is orthogonal. In general,

$$|A(z - X)| \neq |z - X|.$$

Counterexample (one dimension). Let $X \sim N(0,1)$, $z = 0$, and $A = [2]$:

$$D^{L_2}(0 \mid 2X) = (1 + 2\mathbb{E}|X|)^{-1} \neq (1 + \mathbb{E}|X|)^{-1}.$$

Conclusion.

$$\boxed{D^{L_2}(Az \mid AX) \neq D^{L_2}(z \mid X) \text{ for general invertible } A.}$$

1.5.3 L_2-Depth and Axiom D3 (Null at Infinity)

Axiom D3: Formal statement

A depth function satisfies **Null at Infinity** if $|z| \to \infty$ implies $D(z \mid X) \to 0$.

Verification for the L_2-depth

Assume $\mathbb{E}|X| < \infty$. By the reverse triangle inequality,

$$|z - X| \geq |z| - |X|.$$

Taking expectations,

$$\mathbb{E}|z - X| \geq |z| - \mathbb{E}|X| \xrightarrow{|z|\to\infty} \infty,$$

hence

$$D^{L_2}(z \mid X) = \frac{1}{1 + \mathbb{E}|z - X|} \longrightarrow 0.$$

1.5.4 L_2-Depth and Axiom D4 (Monotonicity on Rays)

Axiom D4: Formal statement

Let z^* be a deepest point. For all z and $t \in [0, 1]$,

$$D(z^* + t(z - z^*) \mid X) \geq D(z \mid X).$$

Verification for the L_2-depth

Deepest points minimize $\mathbb{E}|z - X|$ (spatial medians). Convexity of $z \mapsto |z - x|$ implies

$$\mathbb{E}|z^* + t(z - z^*) - X| \leq \mathbb{E}|z - X|,$$

and since $u \mapsto (1 + u)^{-1}$ is decreasing, D4 follows.

1.5.5 L_2-Depth and Axiom D5 (Upper Semicontinuity)

Axiom D5: Formal statement

For every α, the upper level set

$$D_\alpha(X) = \{z : D(z \mid X) \geq \alpha\}$$

is closed.

Verification for the L_2-depth

For $\alpha \in (0, 1]$,

$$D^{L_2}(z \mid X) \geq \alpha \iff \mathbb{E}|z - X| \leq \frac{1}{\alpha} - 1.$$

The function $f(z) = \mathbb{E}|z - X|$ is continuous by dominated convergence, hence its sublevel sets are closed. Therefore L_2-depth is upper semicontinuous.

Final status of L_2-depth

Axiom	Status
D1 (Translation invariance)	✓
D2 (Linear invariance)	×
D3 (Null at infinity)	✓
D4 (Monotonicity on rays)	✓
D5 (Upper semicontinuity)	✓

1.5.6 Affine-Invariant L_2-Depth

Depth under study

Let $X \sim P$ be an $\mathbb{R}^d$-valued random vector. Let $S_X \in \mathbb{R}^{d\times d}$ be a **positive definite, affine-equivariant scatter functional** satisfying

$$S_{AX+b} = AS_XA^\top \quad \text{for all invertible } A \text{ and all } b \in \mathbb{R}^d.$$

Define the Mahalanobis-type norm

$$|u|_{S_X} := \sqrt{u^\top S_X^{-1} u}.$$

The affine-invariant L_2-depth is

$$D_{\text{aff}}^{L_2}(z \mid X) = \left(1 + \mathbb{E}|z - X|_{S_X}\right)^{-1}.$$

Assume throughout that S_X is positive definite and $\mathbb{E}|X|_{S_X} < \infty$.

Axiom D1 — Translation invariance

For all $z, b \in \mathbb{R}^d$,

$$D(z \mid X) = D(z + b \mid X + b).$$

Proof. Let $X' = X + b$ and $z' = z + b$. By affine equivariance,

$$S_{X'} = S_{X+b} = S_X.$$

Hence,

$$|z' - X'|_{S_{X'}} = |(z + b) - (X + b)|_{S_X} = |z - X|_{S_X}.$$

Taking expectations yields

$$D_{\text{aff}}^{L_2}(z + b \mid X + b) = D_{\text{aff}}^{L_2}(z \mid X).$$

Conclusion. Affine-invariant L_2-depth satisfies D1.

Axiom D2 — Linear (Affine) invariance

For every invertible A,

$$D(z \mid X) = D(Az \mid AX).$$

Proof. Let $X' = AX$ and $z' = Az$. By affine equivariance,

$$S_{X'} = S_{AX} = AS_X A^\top.$$

Then

$$|z' - X'|_{S_{X'}} = \sqrt{(Az - AX)^\top (AS_X A^\top)^{-1}(Az - AX)}.$$

Using $(AS_X A^\top)^{-1} = A^{-\top} S_X^{-1} A^{-1}$,

$$|z' - X'|_{S_{X'}} = \sqrt{(z - X)^\top S_X^{-1}(z - X)} = |z - X|_{S_X}.$$

Thus expectations coincide and

$$D^{L_2}_{\text{aff}}(Az \mid AX) = D^{L_2}_{\text{aff}}(z \mid X).$$

Conclusion. Affine-invariant L_2-depth satisfies D2.

Axiom D3 — Null at infinity

By the reverse triangle inequality in Mahalanobis norm,

$$|z - X|_{S_X} \geq |z|_{S_X} - |X|_{S_X}.$$

Taking expectations,

$$\mathbb{E}|z - X|_{S_X} \geq |z|_{S_X} - \mathbb{E}|X|_{S_X}.$$

Since $|z|_{S_X} \to \infty$ as $|z| \to \infty$, it follows that

$$D^{L_2}_{\text{aff}}(z \mid X) \to 0.$$

Conclusion. Affine-invariant L_2-depth satisfies D3.

Axiom D4 — Monotonicity on rays

The argument mirrors the L_2-depth case:

- deepest points z^* minimize $\mathbb{E}|z - X|_{S_X}$;
- $|z - X|_{S_X}$ is convex in z;
- expectation preserves convexity.

Hence, for all $t \in [0, 1]$,

$$D^{L_2}_{\text{aff}}(z^* + t(z - z^*) \mid X) \geq D^{L_2}_{\text{aff}}(z \mid X).$$

Conclusion. Affine-invariant L_2-depth satisfies D4.

Axiom D5 — Upper semicontinuity

The map $z \mapsto |z - X|_{S_X}$ is continuous. Under the finite first-moment assumption, dominated convergence implies continuity of $\mathbb{E}|z - X|_{S_X}$. Therefore, all upper level sets are closed.

Conclusion. Affine-invariant L_2-depth satisfies D5.

Final status — Affine-invariant L_2-depth

Axiom	Status
D1 (Translation invariance)	✓
D2 (Linear invariance)	✓
D3 (Null at infinity)	✓
D4 (Monotonicity on rays)	✓
D5 (Upper semicontinuity)	✓

1.5.7 Mahalanobis Depth

Depth under study

Let $X \sim P$ be an $\mathbb{R}^d$-valued random vector. Let

- $c_X \in \mathbb{R}^d$ be an **affine-equivariant location functional**, satisfying

$$c_{AX+b} = Ac_X + b \quad \text{for all invertible } A \text{ and } b \in \mathbb{R}^d;$$

- $S_X \in \mathbb{R}^{d\times d}$ be a **positive definite affine-equivariant scatter functional**, satisfying

$$S_{AX+b} = AS_XA^\top.$$

The **Mahalanobis depth** is defined as

$$D^{\text{Mah}}(z \mid X) = \left(1 + (z - c_X)^\top S_X^{-1}(z - c_X)\right)^{-1}.$$

Assume throughout that S_X is positive definite.

Axiom D1 — Translation invariance

For all $z, b \in \mathbb{R}^d$,

$$D(z \mid X) = D(z + b \mid X + b).$$

Verification. Let $X' = X + b$ and $z' = z + b$. By affine equivariance,

$$c_{X'} = c_X + b, \qquad S_{X'} = S_X.$$

Hence

$$z' - c_{X'} = z - c_X,$$

and therefore

$$(z' - c_{X'})^\top S_{X'}^{-1}(z' - c_{X'}) = (z - c_X)^\top S_X^{-1}(z - c_X).$$

Substitution yields

$$D^{\mathrm{Mah}}(z + b \mid X + b) = D^{\mathrm{Mah}}(z \mid X).$$

Conclusion. Mahalanobis depth satisfies D1.

Axiom D2 — Linear (Affine) invariance

For all invertible A,

$$D(z \mid X) = D(Az \mid AX).$$

Verification. Let $X' = AX$ and $z' = Az$. By affine equivariance,

$$c_{X'} = Ac_X, \qquad S_{X'} = AS_X A^\top.$$

Then

$$z' - c_{X'} = A(z - c_X).$$

Using $(AS_X A^\top)^{-1} = A^{-\top} S_X^{-1} A^{-1}$,

$$(z' - c_{X'})^\top S_{X'}^{-1}(z' - c_{X'}) = (z - c_X)^\top S_X^{-1}(z - c_X).$$

Hence

$$D^{\mathrm{Mah}}(Az \mid AX) = D^{\mathrm{Mah}}(z \mid X).$$

Conclusion. Mahalanobis depth satisfies D2.

Axiom D3 — Null at infinity

Since S_X^{-1} is positive definite, there exists $\lambda_{\min} > 0$ such that

$$(z - c_X)^\top S_X^{-1}(z - c_X) \geq \lambda_{\min}|z - c_X|^2.$$

As $|z| \to \infty$, also $|z - c_X| \to \infty$, and the quadratic form diverges. Therefore,

$$D^{\mathrm{Mah}}(z \mid X) \to 0.$$

Conclusion. Mahalanobis depth satisfies D3.

Axiom D4 — Monotonicity on rays

The depth is maximized when the quadratic form is minimized. The unique minimizer is

$$z^* = c_X.$$

Let $z_t = c_X + t(z - c_X)$ for $t \in [0, 1]$. Then

$$(z_t - c_X)^\top S_X^{-1}(z_t - c_X) = t^2(z - c_X)^\top S_X^{-1}(z - c_X).$$

Since $t^2 \leq 1$, the quadratic form decreases along the ray toward c_X. As depth is a decreasing function of this form,

$$D^{\text{Mah}}(z_t \mid X) \geq D^{\text{Mah}}(z \mid X).$$

Conclusion. Mahalanobis depth satisfies D4.

Axiom D5 — Upper semicontinuity

The mapping

$$z \mapsto (z - c_X)^\top S_X^{-1}(z - c_X)$$

is a continuous polynomial function. Hence

$$D^{\text{Mah}}(z \mid X) = (1 + \text{quadratic form})^{-1}$$

is continuous, and all its upper level sets are closed.

Conclusion. Mahalanobis depth satisfies D5.

Final status — Mahalanobis depth

Axiom	**Status**
D1 (Translation invariance)	✓
D2 (Linear invariance)	✓
D3 (Null at infinity)	✓
D4 (Monotonicity on rays)	✓
D5 (Upper semicontinuity)	✓

1.5.8 Projection Depth

Depth under study

Let $X \sim P$ be an $\mathbb{R}^d$-valued random vector, and let

$$S^{d-1} = \{p \in \mathbb{R}^d : \ |p| = 1\}.$$

For a real-valued random variable U:

- $\mathrm{med}(U)$ denotes a median;
- $\mathrm{Dmed}(U) := \mathrm{med}(|U - \mathrm{med}(U)|)$ denotes the MAD.

Assume $\mathrm{Dmed}(\langle p, X\rangle) > 0$ for all $p \in S^{d-1}$.

The **projection depth** is defined by

$$D^{\mathrm{proj}}(z \mid X) = \Big(1 + \sup_{p \in S^{d-1}} \frac{|\langle p, z\rangle - \mathrm{med}(\langle p, X\rangle)|}{\mathrm{Dmed}(\langle p, X\rangle)}\Big)^{-1}.$$

Axiom D1 — Translation invariance

Let $b \in \mathbb{R}^d$ and define $z' = z + b$, $X' = X + b$. For any $p \in S^{d-1}$,

$$\langle p, z'\rangle - \mathrm{med}(\langle p, X'\rangle) = \langle p, z\rangle - \mathrm{med}(\langle p, X\rangle).$$

Moreover, $\mathrm{Dmed}(\langle p, X'\rangle) = \mathrm{Dmed}(\langle p, X\rangle)$. Hence the standardized deviations coincide for all p, and

$$D^{\mathrm{proj}}(z + b \mid X + b) = D^{\mathrm{proj}}(z \mid X).$$

Conclusion. Projection depth satisfies D1.

Axiom D2 — Linear (Affine) invariance

Let A be invertible and define $z' = Az$, $X' = AX$. For $p \in S^{d-1}$,

$$\langle p, Az\rangle = \langle A^\top p, z\rangle.$$

Define $q := \frac{A^\top p}{|A^\top p|} \in S^{d-1}$. Then

$$\langle p, Az\rangle = |A^\top p| \langle q, z\rangle, \qquad \langle p, AX\rangle = |A^\top p| \langle q, X\rangle.$$

Using the scaling properties of the median and MAD,

$$\frac{|\langle p, Az\rangle - \mathrm{med}(\langle p, AX\rangle)|}{\mathrm{Dmed}(\langle p, AX\rangle)} = \frac{|\langle q, z\rangle - \mathrm{med}(\langle q, X\rangle)|}{\mathrm{Dmed}(\langle q, X\rangle)}.$$

As p ranges over S^{d-1}, so does q. Therefore,

$$D^{\mathrm{proj}}(Az \mid AX) = D^{\mathrm{proj}}(z \mid X).$$

Conclusion. Projection depth satisfies D2.

Axiom D3 — Null at infinity

Fix any $p \in S^{d-1}$. As $|z| \to \infty$, we have $|\langle p, z\rangle| \to \infty$. Since $\mathrm{med}(\langle p, X\rangle)$ and $\mathrm{Dmed}(\langle p, X\rangle)$ are finite and positive, the standardized deviation diverges. Hence,

$$D^{\mathrm{proj}}(z \mid X) \to 0.$$

Conclusion. Projection depth satisfies D3.

Axiom D4 — Monotonicity on rays

Define the outlyingness function

$$O(z) := \sup_{p \in S^{d-1}} \frac{|\langle p, z\rangle - \mathrm{med}(\langle p, X\rangle)|}{\mathrm{Dmed}(\langle p, X\rangle)}.$$

For fixed p, the mapping in z is convex. The supremum of convex functions is convex, hence $O(z)$ is convex. If z^* minimizes O, then convexity yields

$$O(z^* + t(z - z^*)) \le O(z) \quad \forall\, t \in [0, 1].$$

Since $D^{\mathrm{proj}} = (1 + O)^{-1}$ is decreasing in O,

$$D^{\mathrm{proj}}(z^* + t(z - z^*) \mid X) \ge D^{\mathrm{proj}}(z \mid X).$$

Conclusion. Projection depth satisfies D4.

Axiom D5 — Upper semicontinuity

For each fixed p, the standardized deviation is continuous in z. The supremum over $p \in S^{d-1}$ of continuous functions is upper semicontinuous. Therefore $O(z)$ is upper semicontinuous, and so is

$$D^{\mathrm{proj}}(z \mid X) = (1 + O(z))^{-1}.$$

Upper level sets are closed.

Conclusion. Projection depth satisfies D5.

Final status — Projection depth

1.5.9 Oja Depth

Depth under study

Let $X \sim P$ be an $\mathbb{R}^d$-valued random vector. Let $X_1, \ldots, X_d$ be i.i.d. copies of X. Denote by $\mathrm{co}(\cdot)$ the convex hull and by $\mathrm{Vol}_d(\cdot)$ the d-dimensional volume.

Axiom	**Status**
D1 (Translation invariance)	✓
D2 (Linear invariance)	✓
D3 (Null at infinity)	✓
D4 (Monotonicity on rays)	✓
D5 (Upper semicontinuity)	✓

Let Σ_X be a positive definite scatter matrix. The **Oja depth** is defined as

$$D^{\mathrm{Oja}}(z \mid X) = \left(1 + \frac{\mathbb{E}[\mathrm{Vol}_d(\mathrm{co}(z, X_1, \dots, X_d))]}{\sqrt{\det \Sigma_X}}\right)^{-1}.$$

Assume that Σ_X exists and is positive definite, and that

$$\mathbb{E}\,\mathrm{Vol}_d(\mathrm{co}(z, X_1, \dots, X_d)) < \infty.$$

Axiom D1 — Translation invariance

Let $b \in \mathbb{R}^d$ and define $z' = z + b$, $X_i' = X_i + b$. By translation invariance of volume,

$$\mathrm{Vol}_d(A + b) = \mathrm{Vol}_d(A) \quad \text{for all } A \subset \mathbb{R}^d.$$

Hence,

$$\mathrm{Vol}_d(\mathrm{co}(z', X_1', \dots, X_d')) = \mathrm{Vol}_d(\mathrm{co}(z, X_1, \dots, X_d)).$$

Since $\Sigma_{X+b} = \Sigma_X$, it follows that

$$D^{\mathrm{Oja}}(z + b \mid X + b) = D^{\mathrm{Oja}}(z \mid X).$$

Conclusion. Oja depth satisfies D1.

Axiom D2 — Linear (Affine) invariance

Let A be invertible and define $z' = Az$, $X_i' = AX_i$. Under linear maps, volume transforms as

$$\mathrm{Vol}_d(AA_0) = |\det A|\,\mathrm{Vol}_d(A_0).$$

Therefore,

$$\mathrm{Vol}_d(\mathrm{co}(z', X_1', \dots, X_d')) = |\det A|\,\mathrm{Vol}_d(\mathrm{co}(z, X_1, \dots, X_d)).$$

By affine equivariance of covariance,

$$\Sigma_{AX} = A\Sigma_X A^\top, \qquad \det \Sigma_{AX} = (\det A)^2 \det \Sigma_X.$$

Thus the normalization by $\sqrt{\det \Sigma_X}$ cancels the factor $|\det A|$, and taking expectations yields

$$D^{\mathrm{Oja}}(Az \mid AX) = D^{\mathrm{Oja}}(z \mid X).$$

Conclusion. Oja depth satisfies D2.

Axiom D3 — Null at infinity

Fix $X_1, \ldots, X_d$. As $|z| \to \infty$, the volume

$$\mathrm{Vol}_d(\mathrm{co}(z, X_1, \ldots, X_d))$$

grows at least linearly in $|z|$, since the height of the simplex diverges. Hence,

$$\mathbb{E}\,\mathrm{Vol}_d(\mathrm{co}(z, X_1, \ldots, X_d)) \longrightarrow \infty,$$

and therefore

$$D^{\mathrm{Oja}}(z \mid X) \to 0 \quad \text{as } |z| \to \infty.$$

Conclusion. Oja depth satisfies D3.

Axiom D4 — Monotonicity on rays

Define the Oja objective

$$O(z) := \mathbb{E}\,\mathrm{Vol}_d(\mathrm{co}(z, X_1, \ldots, X_d)).$$

For fixed $x_1, \ldots, x_d$, the mapping

$$z \mapsto \mathrm{Vol}_d(\mathrm{co}(z, x_1, \ldots, x_d))$$

is convex, since simplex volume is the absolute value of a determinant. Expectation preserves convexity, so $O(z)$ is convex.

If z^* minimizes $O(z)$, then convexity implies

$$O(z^* + t(z - z^*)) \le O(z) \quad \forall\, t \in [0, 1].$$

Since depth is a decreasing function of $O(z)$,

$$D^{\mathrm{Oja}}(z^* + t(z - z^*) \mid X) \ge D^{\mathrm{Oja}}(z \mid X).$$

Conclusion. Oja depth satisfies D4.

Axiom D5 — Upper semicontinuity

For fixed $x_1, \ldots, x_d$, the volume function is continuous in z. Under the assumed integrability, dominated convergence implies continuity of

$$z \mapsto \mathbb{E}\,\mathrm{Vol}_d(\mathrm{co}(z, X_1, \ldots, X_d)).$$

Therefore $D^{\mathrm{Oja}}(z \mid X)$ is continuous and hence upper semicontinuous.

Conclusion. Oja depth satisfies D5.

Final status — Oja depth

Axiom	**Status**
D1 (Translation invariance)	✓
D2 (Linear invariance)	✓
D3 (Null at infinity)	✓
D4 (Monotonicity on rays)	✓
D5 (Upper semicontinuity)	✓

1.5.10 Location (Halfspace / Tukey) Depth

Depth under study

We work in $\mathbb{R}^d$ with $d \geq 1$. Let $X \sim P$ be a Borel probability measure on $\mathbb{R}^d$.

A **closed halfspace** in $\mathbb{R}^d$ is any set of the form

$$H_{u,t} = \{x \in \mathbb{R}^d : u^\top x \leq t\},$$

where $u \in \mathbb{R}^d \setminus \{0\}$ and $t \in \mathbb{R}$.

The **location (halfspace / Tukey) depth** of a point $z \in \mathbb{R}^d$ is defined as

$$D^{\mathrm{loc}}(z \mid X) = \inf\{P(H) : H \text{ is a closed halfspace and } z \in H\}.$$

Axiom D1 — Translation invariance

Let $b \in \mathbb{R}^d$, and define $z' = z + b$, $X' = X + b$. If

$$H = \{x : u^\top x \leq t\}$$

is a closed halfspace containing z, then

$$H' = H + b = \{x : u^\top x \leq t + u^\top b\}$$

is a closed halfspace containing z', and $P'(H') = P(H)$. Taking infima yields

$$D^{\mathrm{loc}}(z \mid X) = D^{\mathrm{loc}}(z + b \mid X + b).$$

Conclusion. Halfspace depth satisfies D1.

Axiom D2 — Linear (Affine) invariance

Let A be invertible. If

$$H = \{x : u^\top x \le t\},$$

then

$$A(H) = \{y : (A^{-\top} u)^\top y \le t\}$$

is again a closed halfspace. Moreover,

$$z \in H \iff Az \in A(H), \qquad P_{AX}(A(H)) = P(H).$$

Taking infima over all halfspaces yields

$$D^{\mathrm{loc}}(z \mid X) = D^{\mathrm{loc}}(Az \mid AX).$$

Conclusion. Halfspace depth satisfies D2.

Axiom D3 — Null at infinity

Fix a direction $u \in S^{d-1}$ and define

$$H_z = \{x : u^\top x \le u^\top z\}.$$

Then $z \in H_z$ and

$$D^{\mathrm{loc}}(z \mid X) \le P(H_z).$$

As $|z| \to \infty$, one can choose the sign of u so that $P(H_z) \to 0$ by tightness of probability measures on $\mathbb{R}^d$. Hence

$$D^{\mathrm{loc}}(z \mid X) \to 0.$$

Conclusion. Halfspace depth satisfies D3.

Axiom D4 — Monotonicity on rays

Let z^* be a deepest point and define

$$z_t = z^* + t(z - z^*), \quad t \in [0, 1].$$

If a closed halfspace H contains z_t, convexity of halfspaces implies that whenever $z \notin H$, also $z_t \notin H$. Hence

$$\{H : z \in H\} \subseteq \{H : z_t \in H\}.$$

Taking infima yields

$$D^{\mathrm{loc}}(z_t \mid X) \geq D^{\mathrm{loc}}(z \mid X).$$

Conclusion. Halfspace depth satisfies D4.

Axiom D5 — Upper semicontinuity

For $\alpha \in [0, 1]$,

$$D_\alpha^{\mathrm{loc}}(X) = \{z : D^{\mathrm{loc}}(z \mid X) \geq \alpha\} = \bigcap_{\substack{H \text{ closed halfspace} \\ P(H) \geq \alpha}} H.$$

This is an arbitrary intersection of closed sets, hence closed.

Conclusion. Halfspace depth satisfies D5.

Final status — Halfspace (Tukey) depth

Axiom	Status
D1 (Translation invariance)	✓
D2 (Linear invariance)	✓
D3 (Null at infinity)	✓
D4 (Monotonicity on rays)	✓
D5 (Upper semicontinuity)	✓

1.5.11 Tukey (Halfspace) Depth — Affine Equivariance

What is being proved

We prove that

$$D^{\mathrm{Tukey}}(z \mid X) = D^{\mathrm{Tukey}}(Az + b \mid AX + b)$$

for every $z \in \mathbb{R}^d$, every invertible matrix A, and every $b \in \mathbb{R}^d$.

Ambient setting

Let

$$X : \Omega \to \mathbb{R}^d$$

be a random vector with distribution P, defined by

$$P(B) = \mathbb{P}(X \in B)$$

for every Borel set $B \subset \mathbb{R}^d$.

Closed halfspaces

A closed halfspace is any set of the form

$$H(u,t) := \{x \in \mathbb{R}^d : u^\top x \leq t\},$$

with $u \neq 0$ and $t \in \mathbb{R}$.

Definition of Tukey depth

$$D^{\text{Tukey}}(z \mid X) = \inf \left\{ P(H) : \begin{array}{l} H \text{ is a closed halfspace} \\ z \in H \end{array} \right\}.$$

Affine transformations

An affine transformation is a map

$$T(x) = Ax + b,$$

with A invertible and $b \in \mathbb{R}^d$. The pushforward distribution of $AX + b$ is

$$P_{AX+b}(B) = P(T^{-1}(B)).$$

Lemma 1 — Affine images of halfspaces

Let

$$H = \{x : u^\top x \leq t\}$$

be a halfspace. Its affine image is

$$T(H) = \{y : y = Ax + b \text{ for some } x \in H\}.$$

Substituting $x = A^{-1}(y - b)$ yields

$$(A^{-\top} u)^\top y \leq t + u^\top A^{-1} b,$$

which is again the defining inequality of a closed halfspace.

Conclusion. Affine transformations map halfspaces bijectively to halfspaces.

Lemma 2 — Probability preservation

For any halfspace H,

$$P_{AX+b}(T(H)) = P(H).$$

Main proof

By definition,

$$D^{\text{Tukey}}(Az + b \mid AX + b) = \inf_{\substack{H' \text{ halfspace} \\ Az+b \in H'}} P_{AX+b}(H').$$

By Lemma 1, every such H' can be written uniquely as $H' = T(H)$ for a halfspace H containing z. By Lemma 2, $P_{AX+b}(H') = P(H)$. Therefore,

$$\inf_{H' \ni Az+b} P_{AX+b}(H') = \inf_{H \ni z} P(H) = D^{\text{Tukey}}(z \mid X).$$

Final conclusion

$$\boxed{D^{\text{Tukey}}(Az + b \mid AX + b) = D^{\text{Tukey}}(z \mid X)}$$

Geometric intuition

Tukey depth depends only on sidedness with respect to hyperplanes and the probability mass on each side. Affine transformations preserve sidedness and probability ordering, merely relabeling coordinates.

Exam-ready takeaway

Tukey depth is affine equivariant because affine maps send halfspaces bijectively to halfspaces and preserve probability mass. Therefore, the infimum over halfspaces containing a point is unchanged.

1.5.12 Simplicial Depth

Ambient setting

We work in $\mathbb{R}^d$ with $d \geq 1$. Let

$$X : \Omega \to \mathbb{R}^d$$

be a random vector with probability distribution P. Let

$$X_1, X_2, \ldots, X_{d+1}$$

be independent copies of X.

Convex hulls and simplices

For points $x_1, \ldots, x_k \in \mathbb{R}^d$, the convex hull is

$$\mathrm{co}(x_1, \ldots, x_k) = \left\{ \sum_{i=1}^{k} \lambda_i x_i : \lambda_i \geq 0,\ \sum_{i=1}^{k} \lambda_i = 1 \right\}.$$

A d-simplex is the convex hull of $d+1$ affinely independent points: a line segment in $d = 1$, a triangle in $d = 2$, and a tetrahedron in $d = 3$.

Definition of simplicial depth

The **simplicial depth** of a point $z \in \mathbb{R}^d$ is

$$D^{\mathrm{sim}}(z \mid X) = \mathbb{P}\big(z \in \mathrm{co}(X_1, \ldots, X_{d+1})\big).$$

It is the probability that a random simplex generated by the data contains z.

Axiom D1 — Translation invariance

For any $b \in \mathbb{R}^d$,

$$z + b \in \mathrm{co}(X_1 + b, \ldots, X_{d+1} + b) \iff z \in \mathrm{co}(X_1, \ldots, X_{d+1}).$$

Hence,

$$D^{\mathrm{sim}}(z + b \mid X + b) = D^{\mathrm{sim}}(z \mid X).$$

Conclusion. Simplicial depth satisfies D1.

Axiom D2 — Linear (Affine) invariance

Let A be invertible. By linearity,

$$\mathrm{co}(AX_1, \ldots, AX_{d+1}) = A\ \mathrm{co}(X_1, \ldots, X_{d+1}).$$

Thus,

$$Az \in \mathrm{co}(AX_1, \ldots, AX_{d+1}) \iff z \in \mathrm{co}(X_1, \ldots, X_{d+1}),$$

and therefore,

$$D^{\mathrm{sim}}(Az \mid AX) = D^{\mathrm{sim}}(z \mid X).$$

Conclusion. Simplicial depth satisfies D2.

Axiom D3 — Null at infinity

By tightness of probability measures on $\mathbb{R}^d$, for any $\varepsilon > 0$ there exists a compact set K such that $P(X \in K) \geq 1 - \varepsilon$. Then

$$\mathbb{P}(X_1, \ldots, X_{d+1} \in K) \geq (1 - \varepsilon)^{d+1}.$$

The convex hull $\mathrm{co}(K)$ is compact. For sufficiently large $|z|$,

$$z \notin \mathrm{co}(K),$$

and hence

$$D^{\mathrm{sim}}(z \mid X) \leq 1 - (1 - \varepsilon)^{d+1}.$$

Letting $\varepsilon \to 0$ yields

$$D^{\mathrm{sim}}(z \mid X) \to 0.$$

Conclusion. Simplicial depth satisfies D3.

Axiom D4 — Monotonicity on rays

Fix a realization $(x_1, \ldots, x_{d+1})$. The simplex $\mathrm{co}(x_1, \ldots, x_{d+1})$ is convex. If both z and a deepest point z^* lie in this simplex, then for any $t \in [0, 1]$,

$$z^* + t(z - z^*) \in \mathrm{co}(x_1, \ldots, x_{d+1}).$$

Thus every simplex containing z also contains the point on the ray toward z^*. Taking probabilities yields

$$D^{\mathrm{sim}}(z^* + t(z - z^*) \mid X) \geq D^{\mathrm{sim}}(z \mid X).$$

This holds under standard continuity assumptions ensuring existence of deepest points.

Conclusion. Simplicial depth satisfies D4 (under regularity conditions).

Axiom D5 — Upper semicontinuity

Define

$$f(z, x_1, \ldots, x_{d+1}) = \mathbf{1}\{z \in \mathrm{co}(x_1, \ldots, x_{d+1})\}.$$

For fixed $(x_1, \ldots, x_{d+1})$, this indicator is upper semicontinuous in z. Taking expectations preserves upper semicontinuity. Hence

$$z \mapsto D^{\mathrm{sim}}(z \mid X)$$

is upper semicontinuous, and its upper level sets are closed.

Conclusion. Simplicial depth satisfies D5.

Final status — Simplicial depth

Axiom	**Status**
D1 (Translation invariance)	✓
D2 (Linear invariance)	✓
D3 (Null at infinity)	✓
D4 (Monotonicity on rays)	✓ (under continuity assumptions)
D5 (Upper semicontinuity)	✓

1.5.13 Consequences of the Depth Axioms

We assume throughout that

$$D(\cdot \mid X) : \mathbb{R}^d \to [0, \infty)$$

is a statistical depth function satisfying axioms D1–D4. Axiom D5 will be invoked only when explicitly stated.

Affine Invariance from D1 and D2

Goal. If a depth satisfies translation invariance (D1) and linear invariance (D2), then it is affine invariant.

Derivation. An affine transformation has the form

$$T(x) = Ax + b,$$

with A invertible and $b \in \mathbb{R}^d$.

By D2 (linear invariance),

$$D(z \mid X) = D(Az \mid AX).$$

By D1 (translation invariance), applied to the transformed pair,

$$D(Az \mid AX) = D(Az + b \mid AX + b).$$

Combining the two equalities yields

$$D(z \mid X) = D(Az + b \mid AX + b).$$

Conclusion. Axioms D1 and D2 together imply affine invariance.

Boundedness of Upper Level Sets from D3

Goal. If a depth satisfies D3 (null at infinity), then all its upper level sets are bounded.

Upper level sets. For $\alpha > 0$, define

$$D_\alpha(X) := \{z \in \mathbb{R}^d : D(z \mid X) \geq \alpha\}.$$

Argument. Assume, for contradiction, that $D_\alpha(X)$ is unbounded. Then there exists a sequence $z_n \in D_\alpha(X)$ such that $|z_n| \to \infty$. By D3,

$$D(z_n \mid X) \to 0.$$

However, by definition of $D_\alpha(X)$,

$$D(z_n \mid X) \geq \alpha > 0 \quad \text{for all } n,$$

which is a contradiction.

Conclusion. Every upper level set $D_\alpha(X)$ is bounded.

Star-Shapedness from D4

Definition (star-shaped set). A set $S \subset \mathbb{R}^d$ is star-shaped about a point x_0 if

$$\forall x \in S,\ \forall t \in [0,1], \quad tx + (1-t)x_0 \in S.$$

Goal. If a depth satisfies D4 (monotonicity on rays), then each upper level set is star-shaped about a deepest point.

Derivation. Let z^* be a deepest point, so

$$D(z^* \mid X) = \sup_z D(z \mid X).$$

Fix $\alpha > 0$ and $z \in D_\alpha(X)$, so $D(z \mid X) \geq \alpha$. By D4,

$$D\big(z^* + t(z - z^*) \mid X\big) \geq D(z \mid X) \geq \alpha \quad \forall t \in [0,1].$$

Hence,

$$z^* + t(z - z^*) \in D_\alpha(X) \quad \forall t \in [0,1].$$

Conclusion. Each upper level set $D_\alpha(X)$ is star-shaped about a deepest point z^*.

The Role of D5

Up to this point, none of the arguments required D5:

- Affine invariance follows from D1 + D2.
- Boundedness of central regions follows from D3.
- Star-shapedness follows from D4.

Axiom D5 (upper semicontinuity) adds topological regularity:

- Upper level sets are closed.
- Maximal depth is attained.
- Deepest points exist as actual points, not merely limits.

Summary of Consequences

Consequence	Axioms Used
Affine invariance	D1 + D2
Bounded central regions	D3
Star-shaped central regions	D4
Closed central regions	D5
Existence of medians	D3 + D5

1.5.14 Proposition 2.1 (Central Symmetry)

This proposition states a fundamental consequence of the depth axioms: if a distribution is perfectly symmetric about a point, then that point must be assigned maximal depth by any depth function satisfying axioms D1–D5. The result is purely axiomatic and does not depend on the construction of a specific depth.

Statement of the Proposition

Let $X \sim P$ be a random vector in $\mathbb{R}^d$, and let $D(\cdot \mid X)$ be a depth function satisfying axioms D1–D5. If P is centrally symmetric about $\theta \in \mathbb{R}^d$, then

$$D(\theta \mid X) = \sup_{z \in \mathbb{R}^d} D(z \mid X).$$

Thus, θ is a deepest point.

Central Symmetry

A distribution P is centrally symmetric about θ if

$$X - \theta \overset{d}{=} \theta - X.$$

Equivalently, for every Borel set $B \subset \mathbb{R}^d$,

$$P(\theta + B) = P(\theta - B).$$

Geometrically, reflecting the distribution through θ leaves it unchanged. No direction carries more probability mass than its opposite.

Proof of Proposition 2.1

Step 1: Reduction by translation invariance (D1). Define $Y := X - \theta$. By translation invariance,

$$D(z \mid X) = D(z - \theta \mid Y).$$

In particular,

$$D(\theta \mid X) = D(0 \mid Y).$$

Hence it suffices to prove that the origin is a deepest point when the distribution is centrally symmetric about 0.

Step 2: Consequence of central symmetry. Since $X \overset{d}{=} -X$, the distribution of X is invariant under reflection. Therefore,

$$D(z \mid X) = D(-z \mid X) \quad \forall z \in \mathbb{R}^d.$$

Step 3: Contradiction hypothesis. Assume that 0 is not a deepest point. Then there exists some $z \neq 0$ such that

$$D(z \mid X) > D(0 \mid X).$$

Step 4: Use of symmetry. By Step 2,

$$D(-z \mid X) = D(z \mid X).$$

Thus both z and $-z$ have depth strictly larger than that of 0.

Step 5: Midpoint argument. The midpoint of z and $-z$ is 0. Both endpoints lie in the same upper level set

$$\{x : D(x \mid X) \geq D(z \mid X)\}.$$

Step 6: Monotonicity on rays (D4). By D4, depth cannot decrease when moving along a segment toward a deepest point. Hence the depth at the midpoint cannot be smaller than the depth at the endpoints:

$$D(0 \mid X) \geq \min\{D(z \mid X), D(-z \mid X)\} = D(z \mid X).$$

Step 7: Contradiction. This contradicts the assumption $D(z \mid X) > D(0 \mid X)$. Therefore the assumption was false.

Step 8: Conclusion. Thus,

$$D(0 \mid X) = \sup_{z \in \mathbb{R}^d} D(z \mid X).$$

Undoing the translation yields

$$D(\theta \mid X) = \sup_{z \in \mathbb{R}^d} D(z \mid X).$$

Use of the Axioms

- D1 (translation invariance): reduction to symmetry about the origin.
- D4 (monotonicity on rays): comparison of endpoint and midpoint depths.
- D5 (upper semicontinuity): guarantees existence of deepest points.
- D2 and D3 are not required for this result.

Interpretation

In a centrally symmetric distribution, any point off-center has an equally deep mirror image. If such points were deeper than the center, their midpoint would have to be even deeper. The midpoint is the center itself. Hence the center must be deepest.

1.5.15 Variants of the Depth Axioms

We work throughout in $\mathbb{R}^d$. Let $D(\cdot \mid X)$ be a statistical depth function.

Part A — D4con (Quasiconcavity / Convexity)

A.1 Motivation. Axiom D4 (monotonicity on rays) guarantees star-shaped central regions. However, star-shaped sets may still be non-convex, exhibiting dents or flat sides.

Many commonly used depth functions (Tukey, projection, Mahalanobis) have convex central regions. This motivates a stronger axiom.

A.2 Definition of D4con. A depth function satisfies **D4con** if, for every α,

$$D_\alpha(X) := \{z : D(z \mid X) \geq \alpha\}$$

is a **convex set**.

Reminder: Convex set. A set $C \subset \mathbb{R}^d$ is convex if

$$\forall x, y \in C,\ \forall t \in [0,1], \quad tx + (1-t)y \in C.$$

A.3 Convexity implies star-shapedness. If a set C is convex and contains a point x_0, then it is star-shaped about x_0.

Indeed, for any $x \in C$ and any $t \in [0, 1]$,

$$tx + (1 - t)x_0 \in C,$$

which is exactly the definition of star-shapedness.

$$\boxed{\text{Convex} \;\Rightarrow\; \text{star-shaped}}$$

Therefore,

$$\text{D4con} \;\Rightarrow\; \text{D4},$$

but not conversely.

A.4 Which depths satisfy D4con?

- **Tukey (halfspace) depth**: central regions are intersections of halfspaces.
- **Projection depth**: upper level sets are intersections of slabs.
- **Mahalanobis depth**: upper level sets are ellipsoids.
- **Affine-invariant L_2-depth**: sublevel sets of convex functions.

Depths that generally do **not** satisfy D4con include simplicial depth (under discreteness) and Oja depth (typically star-shaped but not convex).

Convexity yields cleaner geometry, easier optimization, and stronger theoretical control, at the cost of flexibility.

Part B — D2iso (Isometric Invariance)

B.1 Definition of an isometry. A map $T : \mathbb{R}^d \to \mathbb{R}^d$ is an **isometry** if

$$|T(x) - T(y)| = |x - y| \quad \forall x, y.$$

In Euclidean space, every isometry has the form

$$T(x) = Qx + b,$$

where Q is orthogonal and $b \in \mathbb{R}^d$.

B.2 Definition of D2iso. A depth satisfies **D2iso** if

$$D(z \mid X) = D(Qz + b \mid QX + b)$$

for every orthogonal matrix Q and vector b.

B.3 What D2iso preserves.

- Distances

- Angles
- Volumes
- Euclidean geometry

It does **not** preserve general affine structure, anisotropy, or relative scaling.

B.4 Depths and D2iso. Depths satisfying D2iso include L_2-depth, spatial median depth, and Euclidean distance–based depths.

Depths such as Mahalanobis, Tukey, and projection depth generally fail D2iso but satisfy full affine invariance.

B.5 D2iso versus D2. Every isometry is affine, but not every affine transformation is an isometry. Hence,

$$\text{D2} \;\Rightarrow\; \text{D2iso}, \quad \text{but not conversely.}$$

Part C — D2sca (Scale Invariance)

C.1 Definition. A depth satisfies **D2sca** if

$$D(z \mid X) = D(cz \mid cX) \quad \forall c > 0.$$

This allows global expansion and contraction, but not rotation, shear, or anisotropic scaling.

C.2 Hierarchy of invariances. There is a strict hierarchy:

$$\text{Affine invariance (D2)} \;\Rightarrow\; \text{Isometric invariance (D2iso)} \;\Rightarrow\; \text{Scale invariance (D2sca)}.$$

None of the reverse implications hold.

C.3 Examples. Depths satisfying only D2sca include certain univariate rank-based and radial depths.

Depths such as raw L_2-depth and moment Mahalanobis depth (without normalization) fail D2sca.

C.4 Limitations of scale invariance. Without full affine invariance, units of measurement matter, anisotropy is ignored, and multivariate geometry is distorted. Scale invariance is therefore mainly appropriate in univariate settings.

Part D — What is gained and what is lost?

Variant	**What is gained**	**What is lost**
D4con	Convex regions, clean geometry	Flexibility
D2iso	Metric faithfulness	Affine robustness
D2sca	Unit invariance	Directional information