Classification with mixtures of curved Mahalanobis metrics
— or LMNN in Cayley-Klein geometries —
arXiv:1609.07082

Frank Nielsen¹,²  Boris Muzellec¹  Richard Nock³,⁴,⁵

¹Ecole Polytechnique, France ²Sony CSL, Japan ³Data61, Australia ⁴ANU, Australia ⁵The University of Sydney, Australia

26th September 2016
Mahalanobis distances

- For $Q \succ 0$, a symmetric positive definite matrix like a covariance matrix, define **Mahalanobis distance**:

$$D_Q(p, q) = \sqrt{(p - q)^\top Q (p - q)}$$

Metric distance (indiscernibles/symmetry/triangle inequality)
Eg., $Q = \text{precision matrix } \Sigma^{-1}$, where $\Sigma = \text{covariance matrix}$

- Generalize Euclidean distance when $Q = I$: $D_I(p, q) = \|p - q\|$

- Mahalanobis distance interpreted as Euclidean distance after *Cholesky decomposition* $Q = L^\top L$ and **affine transformation** $x' \leftarrow L^\top x$:

$$D_Q(p, q) = D_I(L^\top p, L^\top q) = \|p' - q'\|$$
Generalizing Mahalanobis distances with Cayley-Klein projective geometries + Learning in Cayley-Klein spaces
Cayley-Klein geometry: Projective geometry [7, 3]

- $\mathbb{RP}^d$: $(\lambda x, \lambda) \sim (x, 1)$
  
  homogeneous coordinates $x \mapsto \tilde{x} = (x, w = 1)$, and
  
  dehomogeneization by “perspective division” $\tilde{x} \mapsto \frac{x}{w}$

- **cross-ratio** measure is *invariant* by
  
  projectivity/homography/collineation:

  $$(p, q; P, Q) = \frac{(p - P)(q - Q)}{(p - Q)(q - P)}$$

  where $p, q, P, Q$ are collinear
Definition of Cayley-Klein geometries

A Cayley-Klein geometry is $\mathcal{K} = (\mathcal{F}, c_{\text{dist}}, c_{\text{angle}})$:

1. A fundamental conic: $\mathcal{F}$

2. A constant unit $c_{\text{dist}} \in \mathbb{C}$ for measuring distances

3. A constant unit $c_{\text{angle}} \in \mathbb{C}$ for measuring angles

See monograph [7]
Distance in Cayley-Klein geometries

\[ \text{dist}(p, q) = c_{\text{dist}} \log((p, q; P, Q)) \]

where \( P \) and \( Q \) are intersection points of line \( l = (pq) \) (\( \tilde{l} = \tilde{p} \times \tilde{q} \) in 2D) with the conic.
Log is principal complex logarithm (modulo \( 2\pi i \))
Key properties of Cayley-Klein distances

- \( \text{dist}(p, p) = 0 \) (law of indiscernibles)

- **Signed** distances: \( \text{dist}(p, q) = -\text{dist}(q, p) \)

- When \( p, q, r \) are collinear

  \[
  \text{dist}(p, q) = \text{dist}(p, r) + \text{dist}(r, q)
  \]

  Geodesics in Cayley-Klein geometries are **straight lines**
  (eventually clipped within the conic domain)

Logarithm is transferring **multiplicative properties** of the cross-ratio to **additive properties** of Cayley-Klein distances.

When \( p, q, P, Q \) are collinear:

\[
(p, q; P, Q) = (p, r; P, Q) \cdot (r, q; P, Q)
\]
Dual conics

In projective geometry, *points* and *lines* are dual concepts.

Dual parameterizations of the fundamental conic $\mathcal{F} = (A, A^\Delta)$

Quadratic form $Q_A(x) = \tilde{x}^\top A\tilde{x}$

- primal conic = set of border points: $\mathcal{C}_A = \{\tilde{p} : Q_A(\tilde{p}) = 0\}$

- dual conic = set of tangent hyperplanes:
  $\mathcal{C}_A^* = \{\tilde{l} : Q_{A^\Delta}(\tilde{l}) = 0\}$

$A^\Delta = A^{-1}|A|$ is the **adjoint matrix**

Adjoint can be computed even when $A$ is not invertible ($|A| = 0$)
Taxonomy

Signature of matrix = sign of eigenvalues of its eigen decomposition

<table>
<thead>
<tr>
<th>Type</th>
<th>$A$</th>
<th>$A^A$</th>
<th>Conic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elliptic</td>
<td>$(+ , + , +)$</td>
<td>$(+ , + , +)$</td>
<td>non-degenerate complex conic</td>
</tr>
<tr>
<td>Hyperbolic</td>
<td>$(+ , + , -)$</td>
<td>$(+ , + , -)$</td>
<td>non-degenerate real conic</td>
</tr>
<tr>
<td>Dual Euclidean</td>
<td>$(+ , 0 , 0)$</td>
<td>$(+ , 0 , 0)$</td>
<td>Two complex lines with a real intersection point</td>
</tr>
<tr>
<td>Dual Pseudo-euclidean</td>
<td>$(+ , - , 0)$</td>
<td>$(+ , 0 , 0)$</td>
<td>Two real lines with a double real intersection point Deux</td>
</tr>
<tr>
<td>Euclidean</td>
<td>$(+ , 0 , 0)$</td>
<td>$(+ , + , 0)$</td>
<td>Two complex points with a double real line passing through</td>
</tr>
<tr>
<td>Pseudo-euclidean</td>
<td>$(+ , 0 , 0)$</td>
<td>$(+ , - , 0)$</td>
<td>Two complex points with a double real line passing through</td>
</tr>
<tr>
<td>Galilean</td>
<td>$(+ , 0 , 0)$</td>
<td>$(+ , 0 , 0)$</td>
<td>Double real line with a real intersection point</td>
</tr>
</tbody>
</table>

Degenerate cases are obtained as limit of non-degenerate cases.

Measurements can be elliptic, hyperbolic or parabolic (degenerate case).
Real CK distances without cross-ratio expressions

For **real** Cayley-Klein measures, we choose the constants:

- **Constants** ($κ$ is curvature):
  - Elliptic ($κ > 0$): $c_{\text{dist}} = \frac{κ}{2i}$
  - Hyperbolic ($κ < 0$): $c_{\text{dist}} = -\frac{κ}{2}$

- Bilinear form $S_{pq} = (p^\top, 1)^\top S(q, 1) = \tilde{p}^\top S\tilde{q}$

- Get rid of **cross-ratio** using:

\[
(p, q; P, Q) = \frac{S_{pq} + \sqrt{S_{pq}^2 - S_{pp}S_{qq}}}{S_{pq} - \sqrt{S_{pq}^2 - S_{pp}S_{qq}}}
\]
Elliptic Cayley-Klein metric distance

\[ d_E(p, q) = \frac{\kappa}{2i} \log \left( \frac{S_{pq} + \sqrt{S_{pq}^2 - S_{pp}S_{qq}}}{S_{pq} - \sqrt{S_{pq}^2 - S_{pp}S_{qq}}} \right) \]

\[ d_E(p, q) = \kappa \arccos \left( \frac{S_{pq}}{\sqrt{S_{pp}S_{qq}}} \right) \]

Notice that \( d_E(p, q) < \kappa \pi \), domain \( \mathbb{D}_S = \mathbb{R}^d \) in elliptic case.

Gnomonic projection \( d_E(x, y) = \kappa \cdot \arccos (\langle x', y' \rangle) \)
Hyperbolic Cayley-Klein distance

When \( p, q \in \mathbb{D}_S := \{ p : S_{pp} < 0 \} \), the hyperbolic domain:

\[
d_H(p, q) = -\frac{\kappa}{2} \log \left( \frac{S_{pq} + \sqrt{S_{pq}^2 - S_{pp}S_{qq}}}{S_{pq} - \sqrt{S_{pq}^2 - S_{pp}S_{qq}}} \right)
\]

\[
d_H(p, q) = -\kappa \arctanh \left( \sqrt{1 - \frac{S_{pp}S_{qq}}{S_{pq}^2}} \right)
\]

\[
d_H(p, q) = -\kappa \arccosh \left( \frac{S_{pq}}{\sqrt{S_{pp}S_{qq}}} \right)
\]

with \( \arccosh(x) = \log(x + \sqrt{x^2 - 1}) \) and \( \arctanh(x) = \frac{1}{2} \log \frac{1+x}{1-x} \). Curvature \( \kappa < 0 \)
Decomposition of the bilinear form [1]

Write \( S = \begin{bmatrix} \Sigma & a \\ a^\top & b \end{bmatrix} = S_{\Sigma,a,b} \) with \( \Sigma \succ 0 \).

\[ S_{p,q} = \tilde{p}^\top S \tilde{q} = p^\top \Sigma q + p^\top a + a^\top q + b \]

Let \( \mu = -\Sigma^{-1}a \in \mathbb{R}^d \) (\( a = -\Sigma \mu \)) and \( b = \mu^\top \Sigma \mu + \text{sign}(\kappa) \frac{1}{\kappa^2} \)

\[ \kappa = \begin{cases} 
(b - \mu^\top \mu)^{-\frac{1}{2}} & b > \mu^\top \mu \\
-(\mu^\top \mu - b)^{-\frac{1}{2}} & b < \mu^\top \mu
\end{cases} \]

Then the bilinear form writes as:

\[ S(p, q) = S_{\Sigma,\mu,\kappa}(p, q) = (p - \mu)^\top \Sigma (q - \mu) + \text{sign}(\kappa) \frac{1}{\kappa^2} \]
Curved Mahalanobis metric distances

We have [1]:

\[
\lim_{\kappa \to 0^+} D_{\Sigma,\mu,\kappa}(p, q) = \lim_{\kappa \to 0^-} D_{\Sigma,\mu,\kappa}(p, q) = D_{\Sigma}(p, q)
\]

Mahalanobis distance \( D_{\Sigma}(p, q) = D_{\Sigma,0,0}(p, q) \)

Thus hyperbolic/elliptic Cayley-Klein distances can be interpreted as **curved Mahalanobis distances**, or \( \kappa \)-Mahalanobis distances

When \( S = \text{diag}(1, 1, \ldots, 1, -1) \), we recover the canonical hyperbolic distance [5] in Cayley-Klein model:

\[
D_h(p, q) = \text{arccosh} \left( \frac{1 - \langle p, q \rangle}{\sqrt{1 - \langle p, p \rangle} \sqrt{1 - \langle q, q \rangle}} \right)
\]

defined inside the interior of a unit ball.
Cayley-Klein bisectors are affine

Bisector $\text{Bi}(p, q)$:

$$\text{Bi}(p, q) = \{ x \in \mathbb{D}_S : \text{dist}_S(p, x) = \text{dist}_S(x, q) \}$$

$$\frac{S(p, x)}{\sqrt{S(p, p)}} = \frac{S(q, x)}{\sqrt{S(q, q)}}$$

arccos and arccosh are monotonically increasing functions.

$$\left\langle x, \sqrt{|S(p, p)|} \Sigma q - \sqrt{|S(q, q)|} \Sigma p \right\rangle$$

$$+ \sqrt{|S(p, p)|}(a^\top(q + x) + b) - \sqrt{|S(q, q)|}(a^\top(p + x) + b) = 0$$

Hyperplanes (restricted to the domain)
Cayley-Klein Voronoi diagrams are affine

Can be computed from equivalent (clipped) power diagrams [2, 5]
https://www.youtube.com/watch?v=YHq3-RL58
Cayley-Klein balls

Blue: Mahalanobis       Red: elliptic       Green: Hyperbolic

Cayley-Klein balls have Mahalanobis ball shapes with displaced centers
Learning curved Mahalanobis metrics
Large Margin Nearest Neighbors [8], LMNN

Learn Mahalanobis distance $M = L^T L > 0$ for a given input data-set $\mathcal{P}$

- Distance of each point to its target neighbors shrink, $\epsilon_{\text{pull}}(L)$
  $$\mathcal{S} = \{(x_i, x_j) : y_i = y_j \text{ and } x_j \in N(x_j)\}$$

- Keep a distance margin of each point to its impostors, $\epsilon_{\text{push}}(L)$
  $$\mathcal{R} = \{(x_i, x_j, x_l) : (x_i, x_j) \in \mathcal{S} \text{ and } y_i \neq y_l\}$$

LMNN: Cost function and optimization

Objective cost function [8]: convex and piecewise linear (SDP)

\[ \epsilon_{\text{pull}}(L) = \sum_{i,i \rightarrow j} \| L(x_i - x_j) \|^2, \]
\[ \epsilon_{\text{push}}(L) = \sum_{i,i \rightarrow j} \sum_{j} (1 - y_{il}) \left[ 1 + \| L(x_i - x_j) \|^2 - \| L(x_i - x_l) \|^2 \right]_+, \]
\[ \epsilon(L) = (1 - \mu) \epsilon_{\text{pull}}(L) + \mu \epsilon_{\text{push}}(L) \]

\( i \rightarrow j \): \( x_j \) is a target neighbor of \( x_i \)
\( y_{il} = 1 \) iff \( x_i \) and \( x_j \) have same label, \( y_{il} = 0 \) otherwise.
\( \mu \) set by cross-validation

Optimize by gradient descent: \( \epsilon(L_{t+1}) = \epsilon(L_t) - \gamma \frac{\partial \epsilon(L_t)}{\partial L} \)

\[ \frac{\partial \epsilon}{\partial L} = (1 - \mu) \sum_{i,i \rightarrow j} C_{ij} + \mu \sum_{(i,j,l) \in \mathcal{R}_t} (C_{ij} - C_{il}) \]

where \( C_{ij} = (x_i - x_j)^\top (x_i - x_j) \)

Easy, no projection mechanism like for Mahalanobis Metric for Clustering (MMC) [9]
Elliptic Cayley-Klein LMNN [1], CVPR 2015

\[
\epsilon(L) = (1 - \mu) \sum_{i,i \to j} d_E(x_i, x_j) + \mu \sum_{i,i \to j} \sum_{l} (1 - y_{il}) \zeta_{ijl}
\]

with \( \zeta_{ijl} = [1 + d_E(x_i, x_j) - d_E(x_i, x_l)]_+ \) (hinge loss)

\[
\frac{\partial \epsilon(L)}{\partial L} = (1 - \mu) \sum_{i,i \to j} \frac{\partial d_E(x_i, x_j)}{\partial L} + \mu \sum_{i,i \to j} \sum_{l} (1 - y_{il}) \frac{\partial \zeta_{ijl}}{\partial L}
\]

\[
C_{ij} = (x_i^\top, 1)^\top (x_j^\top, 1)
\]

\[
\frac{\partial d_E(x_i, x_j)}{\partial L} = \frac{k}{\sqrt{S_{ii} S_{jj} - S_{ij}^2}} L \left( \frac{S_{ij}}{S_{ii}} C_{ii} + \frac{S_{ij}}{S_{jj}} C_{jj} - (C_{ij} + C_{ji}) \right)
\]

\[
\frac{\partial \zeta_{ijl}}{\partial L} = \begin{cases} 
\frac{\partial d_E(x_i, x_j)}{\partial L} - \frac{\partial d_E(x_i, x_l)}{\partial L}, & \text{if } \zeta_{ijl} \geq 0, \\
0, & \text{otherwise.}
\end{cases}
\]
Hyperbolic Cayley-Klein LMNN (new case)

To ensure $S$ keeps correct signature $(1, d, 0)$ during the LMNN gradient descent, we decompose $S = L^\top DL$ (with $L \succ 0$) and perform a gradient descent on $L$ with the following gradient:

$$
\frac{\partial d_H(x_i, x_j)}{\partial L} = \frac{k}{\sqrt{S_{ij}^2 - S_{ii}S_{jj}}} DL \left( \frac{S_{ij}}{S_{ii}} C_{ii} + \frac{S_{ij}}{S_{jj}} C_{jj} - (C_{ij} + C_{ji}) \right)
$$

Recall two difficulties of hyperbolic case compared to elliptic case:

- Hyperbolic Cayley-Klein distance may be very large (unbounded vs. $< \kappa \pi$ for elliptic case)

- Data-set should be contained inside the compact domain $\mathbb{D}_S$
HCK-LMNN: Initialization and learning rate

- Initialize $L = \begin{pmatrix} L' \\ 1 \end{pmatrix}$ and $D$ so that $\mathcal{P} \in \mathbb{D}_S$ with $\Sigma^{-1} = L'\top L'$ (eg., precision matrix of $\mathcal{P}$).

$$D = \begin{pmatrix} -1 & & \\ & \ddots & \\ & & -1 \end{pmatrix} + \kappa \max_x \|L'x\|^2$$

with $\kappa > 1$.

- At iteration $t$, it may happen that $\mathcal{P} \not\in \mathbb{D}_{S_t}$ since we do not
Curved Mahalanobis learning: Results

Experimental results on some UCI data-sets

<table>
<thead>
<tr>
<th>k</th>
<th>Data-set</th>
<th>elliptic</th>
<th>Hyperbolic</th>
<th>Mahalanobis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>wine</td>
<td>0.989</td>
<td>0.865</td>
<td>0.984</td>
</tr>
<tr>
<td></td>
<td>vowel</td>
<td>0.832</td>
<td>0.797</td>
<td>0.827</td>
</tr>
<tr>
<td></td>
<td>balance</td>
<td>0.924</td>
<td>0.891</td>
<td>0.846</td>
</tr>
<tr>
<td></td>
<td>pima</td>
<td>0.726</td>
<td>0.706</td>
<td>0.709</td>
</tr>
<tr>
<td>3</td>
<td>wine</td>
<td>0.983</td>
<td>0.871</td>
<td>0.984</td>
</tr>
<tr>
<td></td>
<td>vowel</td>
<td>0.828</td>
<td>0.782</td>
<td>0.827</td>
</tr>
<tr>
<td></td>
<td>balance</td>
<td>0.917</td>
<td>0.911</td>
<td>0.846</td>
</tr>
<tr>
<td></td>
<td>pima</td>
<td>0.706</td>
<td>0.695</td>
<td>0.709</td>
</tr>
<tr>
<td>5</td>
<td>wine</td>
<td>0.983</td>
<td>0.805</td>
<td>0.984</td>
</tr>
<tr>
<td></td>
<td>vowel</td>
<td>0.826</td>
<td>0.805</td>
<td>0.827</td>
</tr>
<tr>
<td></td>
<td>balance</td>
<td>0.907</td>
<td>0.895</td>
<td>0.846</td>
</tr>
<tr>
<td></td>
<td>pima</td>
<td>0.714</td>
<td>0.712</td>
<td>0.709</td>
</tr>
<tr>
<td>11</td>
<td>wine</td>
<td>0.994</td>
<td>0.983</td>
<td>0.984</td>
</tr>
<tr>
<td></td>
<td>vowel</td>
<td>0.839</td>
<td>0.767</td>
<td>0.827</td>
</tr>
<tr>
<td></td>
<td>balance</td>
<td>0.874</td>
<td>0.897</td>
<td>0.846</td>
</tr>
<tr>
<td></td>
<td>pima</td>
<td>0.713</td>
<td>0.698</td>
<td>0.709</td>
</tr>
</tbody>
</table>

For classification, enough to consider $\kappa \in \{-1, 0, +1\}$
Avoid to compute $d_E$ or $d_H$ for arbitrary $S$

Apply *spectral decomposition* (elliptic case $S = L^\top L$, or hyperbolic case $S = L^\top D L$) and perform coordinate changes so that we consider the canonical metric distances:

$$d_E(x', y') = \arccos \left( \frac{\langle x', y' \rangle}{\|x'\| \|y'\|} \right),$$

$$d_H(x', y') = \arccosh \left( \frac{1 - \langle x', y' \rangle}{\sqrt{1 - \langle x', x' \rangle} \sqrt{1 - \langle y', y' \rangle}} \right)$$

Proximity query: Eg, Vantage Point Tree data-structures [10, 6] (with metric pruning).
Mixed curved Mahalanobis distance

\[ d(x, y) = \alpha d_E(x, y) + (1 - \alpha) d_H(x, y) \]

1. Sum of Riemannian metric distances is metric ("blending" positive with negative constant curvatures)

2. Mixed of bounded distance (elliptic CK) with unbounded distance (hyperbolic CK), hyperparameter tuning \( \alpha \)

<table>
<thead>
<tr>
<th>Datasets</th>
<th>Mahalanobis</th>
<th>elliptic</th>
<th>Hyperbolic</th>
<th>Mixed</th>
<th>( \alpha )</th>
<th>( \beta = (1 - \alpha) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wine</td>
<td>0.993</td>
<td>0.984</td>
<td>0.893</td>
<td>0.986</td>
<td>0.741</td>
<td>0.259</td>
</tr>
<tr>
<td>Sonar</td>
<td>0.733</td>
<td>0.788</td>
<td>0.640</td>
<td>0.802</td>
<td>0.794</td>
<td>0.206</td>
</tr>
<tr>
<td>Balance</td>
<td>0.846</td>
<td>0.910</td>
<td>0.904</td>
<td>0.920</td>
<td>0.440</td>
<td>0.560</td>
</tr>
<tr>
<td>Pima</td>
<td>0.709</td>
<td>0.712</td>
<td>0.699</td>
<td>0.720</td>
<td>0.584</td>
<td>0.416</td>
</tr>
<tr>
<td>Vowel</td>
<td>0.827</td>
<td>0.825</td>
<td>0.816</td>
<td>0.841</td>
<td>0.407</td>
<td>0.593</td>
</tr>
</tbody>
</table>

Although mixed CK distance is a Riemannian metric distance, it is not of constant curvature.
Conclusion
Contributions and perspectives

- Study of Cayley-Klein elliptic/hyperbolic geometries: Affine bisector, Voronoi diagrams from (clipped) power diagrams, Cayley-Klein balls (Mahalanobis shapes with displaced centers), etc.

- Classification with Large Margin Nearest Neighbor (LMNN) in Cayley-Klein elliptic/hyperbolic geometries (hyperbolic geometry: compact domain & unbounded distance)

- Experiments on mixed Cayley-Klein distances

Ongoing work:

Extensions of Cayley-Klein geometries to Machine Learning
Thank you!

https://www.lix.polytechnique.fr/~nielsen/CayleyKlein/

10.1109/ICIP.2016.7532355

arXiv: 1609. 07082
Yanhong Bi, Bin Fan, and Fuchao Wu.
Beyond Mahalanobis metric: Cayley-Klein metric learning.
In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015.

Jean-Daniel Boissonnat, Frank Nielsen, and Richard Nock.
Bregman Voronoi diagrams.

C. Gunn.

Frank Nielsen, Boris Muzellec, and Richard Nock.
Classification with mixtures of curved Mahalanobis metrics.

Frank Nielsen and Richard Nock.
Hyperbolic Voronoi diagrams made easy.

Frank Nielsen, Paolo Piro, and Michel Barlaud.
Bregman vantage point trees for efficient nearest neighbor queries.

Jürgen Richter-Gebert.
*Perspectives on Projective Geometry: A Guided Tour Through Real and Complex Geometry.*

Kilian Q. Weinberger, John Blitzer, and Lawrence K. Saul.
Distance metric learning for large margin nearest neighbor classification.
Eric P. Xing, Andrew Y. Ng, Michael I. Jordan, and Stuart Russell.
Distance metric learning, with application to clustering with side-information.

Peter N. Yianilos.
Data structures and algorithms for nearest neighbor search in general metric spaces.
Overview

Review of Mahalanobis distances

Basics of Cayley-Klein geometry
  - Distance from cross-ratio measures
  - Distance expressions
  - Dual conics

Cayley-Klein distances as curved Mahalanobis distances

Computational geometry in Cayley-Klein geometries

Learning curved Mahalanobis metrics
  - Large Margin Nearest Neighbors (LMNN)
  - Elliptic Cayley-Klein LMNN
  - Hyperbolic Cayley-Klein LMNN
  - Experimental results
  - Nearest-neighbor classification in Cayley-Klein geometries
  - Mixed curved Mahalanobis distance

Contributions and perspectives

Bibliography

Supplemental information
Properties of the cross-ratio

\[
\begin{align*}
\blacktriangleright \quad (p, p; P, P) &= 1 \\
\blacktriangleright \quad (p, q; Q, P) &= \frac{1}{(p, q; P, Q)} \\
\blacktriangleright \quad (p, q; P, Q) &= (p, r; P, Q) \cdot (r, q; P, Q) \text{ when } r \text{ is collinear with } p, q, P, Q
\end{align*}
\]
Measuring angles in Cayley-Klein geometries

\[ \text{angle}(l, m) = c_{\text{angle}} \log((l, m; L, M)) \]

where \( L \) and \( M \) are tangent lines to \( A \) passing through the intersection point \( p \) (\( p = l \times m \) in 2D) of \( l \) \( m \).
Interpretation of hyperbolic Cayley-Klein distance

\[ d_H(x, y) = \kappa \arccosh \left( \prec x', y' \succ \right) \]
Cayley-Klein Voronoi diagrams from (clipped) power diagrams

\[ c_i = \frac{\sum p_i + a}{2\sqrt{S_{p_ip_i}}} \]

\[ r_i^2 = \frac{||\sum p_i + a||^2}{4S_{p_ip_i}} + \frac{a^\top p_i + b}{\sqrt{S_{p[ip_i]}}} \]
Cayley-Klein balls have Mahalanobis ball shapes

Elliptic Cayley-Klein ball case:

\[ \Sigma' = \tilde{r}^2 \Sigma - aa^\top \]
\[ c' = \Sigma'^{-1}(b'a' - \tilde{r}^2 a) \quad \text{with} \quad a' = \Sigma c + a \]
\[ r'^2 = b'^2 - \tilde{r}^2 b + \langle c', c' \rangle_{\Sigma'} \]
\[ b' = a^\top c + b \]
Cayley-Klein balls have Mahalanobis ball shapes

Hyperbolic Cayley-Klein ball case:

\[ \Sigma' = aa^\top - \tilde{r}^2 \Sigma \]
\[ c' = \Sigma'^{-1}(\tilde{r}^2 a - b'a') \]
\[ r'^2 = \tilde{r}^2 b - b'^2 + \langle c', c' \rangle_{\Sigma'} \]

with
\[ a' = \Sigma c + a \]
\[ b' = a^\top c + b \]

... and drawing a Mahalanobis ball amounts to draw a Euclidean ball after affine transformation \( x' \leftarrow L^\top x \).
Spectral decomposition and signature

- **Eigenvalue decomposition:** \( S = O\Lambda O^\top. \)
  \[
  \Lambda = \text{diag}(\Lambda_1, \ldots, \Lambda_{d+1, d+1})
  \]

- **Canonical decomposition:** \( S = OD^{1\frac{1}{2}} \begin{bmatrix} I & 0 \\ 0 & \lambda \end{bmatrix} D^{1\frac{1}{2}}O^\top, \) where \( \lambda \in \{-1, 1\} \) and \( O = \) orthogonal matrix (\( O^{-1} = O^\top \))

- **Diagonal matrix** \( D \) has all positive values, with \( D_{i,i} = \Lambda_{i,i} \) and \( D_{d+1,d+1} = |\Lambda_{d+1,d+1}| \)