Classification with mixtures of curved Mahalanobis metrics

— or LMNN in Cayley-Klein geometries —

arXiv:1609.07082

Frank Nielsen[1,2] Boris Muzellec[1] Richard Nock[3,4,5]

[1]Ecole Polytechnique, France [2]Sony CSL, Japan [3]Data61, Australia [4]ANU, Australia [4]The 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**:

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

  Metric distance (indiscernibles/symmetry/triangle inequality)
  Eg., $Q =$ *precision matrix* Σ^{-1}, where $\Sigma =$ *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** $x' \leftarrow L^\top x$:

$$\boxed{D_Q(p,q) = D_I(L^\top p, L^\top q) = \|p' - q'\|}$$

Generalizing Mahalanobis distances with Cayley-Klein projective geometries + Learning in Cayley-Klein spaces

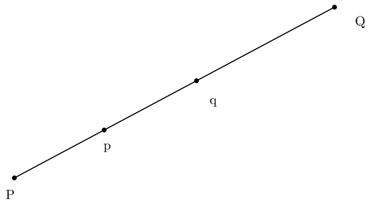
Cayley-Klein geometry: Projective geometry [7, 3]

- $\mathbb{RP}^d$: $(\lambda x, \lambda) \sim (x, 1)$
  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 $\mathcal{K} = (\mathcal{F}, c_{\mathrm{dist}}, c_{\mathrm{angle}})$:

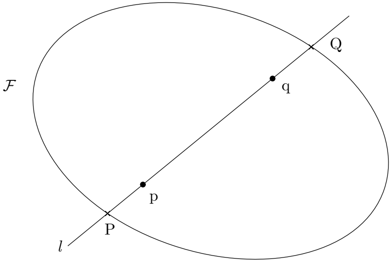
1. A fundamental conic: $\mathcal{F}$
2. A constant unit $c_{\mathrm{dist}} \in \mathbb{C}$ for measuring distances
3. A constant unit $c_{\mathrm{angle}} \in \mathbb{C}$ for measuring angles

See monograph [7]

Distance in Cayley-Klein geometries

$$\boxed{\mathrm{dist}(p, q) = c_{\mathrm{dist}} \ \mathrm{Log}((p, q; P, Q))}$$

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

- $\mathrm{dist}(p, p) = 0$ (law of indiscernibles)
- **Signed** distances : $\mathrm{dist}(p, q) = -\mathrm{dist}(q, p)$
- When p, q, r are collinear

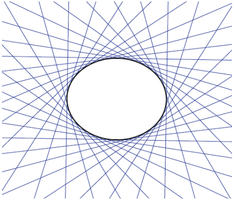
$$\boxed{\mathrm{dist}(p, q) = \mathrm{dist}(p, r) + \mathrm{dist}(r, q)}$$

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

Type	A	A^{Δ}	Conic
Elliptic	$(+,+,+)$	$(+,+,+)$	non-degenerate complex conic
Hyperbolic	$(+,+,-)$	$(+,+,-)$	non-degenerate real conic
Dual Euclidean	$(+,+,0)$	$(+,+,0)$	Two complex lines with a real intersection point
Dual Pseudo-euclidean	$(+,-,0)$	$(+,0,0)$	Two real lines with a double real intersection point Deux
Euclidean	$(+,0,0)$	$(+,+,0)$	Two complex points with a double real line passing through
Pseudo-euclidean	$(+,0,0)$	$(+,-,0)$	Two complex points with a double real line passing through
Galilean	$(+,0,0)$	$(+,0,0)$	Double real line with a real intersection point

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 ($\kappa > 0$): $c_{\mathrm{dist}} = \frac{\kappa}{2i}$
  - Hyperbolic ($\kappa < 0$): $c_{\mathrm{dist}} = -\frac{\kappa}{2}$
- Bilinear form $S_{pq} = (p^\top, 1)^\top S(q, 1) = \tilde{p}^\top S \tilde{q}$
- Get rid of *cross-ratio* using:

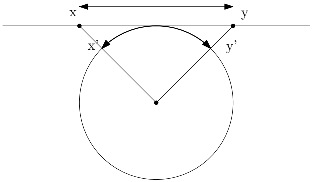
$$(p, q; P, Q) = \frac{S_{pq} + \sqrt{S_{pq}^2 - S_{pp}S_{qq}}}{S_{pq} - \sqrt{S_{pq}^2 - S_{pp}S_{qq}}}$$

Elliptic Cayley-Klein metric distance

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

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

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



Gnomonic projection $d_E(x,y) = \kappa \cdot \arccos\left(\langle x', y' \rangle\right)$

Hyperbolic Cayley-Klein distance

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

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

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

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

with $\mathrm{arccosh}(x) = \log(x + \sqrt{x^2 - 1})$ and $\mathrm{arctanh}(x) = \frac{1}{2}\log\frac{1+x}{1-x}$.
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 $b = \mu^\top \Sigma \mu + \mathrm{sign}(\kappa)\frac{1}{\kappa^2}$

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

Then the bilinear form writes as:

$$\boxed{S(p,q) = S_{\Sigma,\mu,\kappa}(p,q) = (p-\mu)^\top \Sigma (q-\mu) + \mathrm{sign}(\kappa)\frac{1}{\kappa^2}}$$

Curved Mahalanobis metric distances

We have [1]:

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

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 κ-Mahalanobis distances

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

$$D_h(p,q) = \mathrm{arccosh}\left(\frac{1-\langle p,q\rangle}{\sqrt{1-\langle p,p\rangle}\sqrt{1-\langle q,q\rangle}}\right)$$

defined inside the interior of a unit ball.

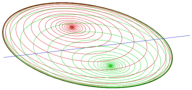
Cayley-Klein bisectors are affine

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

$$\mathrm{Bi}(p,q) = \{x \in \mathbb{D}_S \ :\ \mathrm{dist}_S(p,x) = \mathrm{dist}_S(x,q)\}$$
$$\frac{S(p,x)}{\sqrt{S(p,p)}} = \frac{S(q,x)}{\sqrt{S(q,q)}}$$

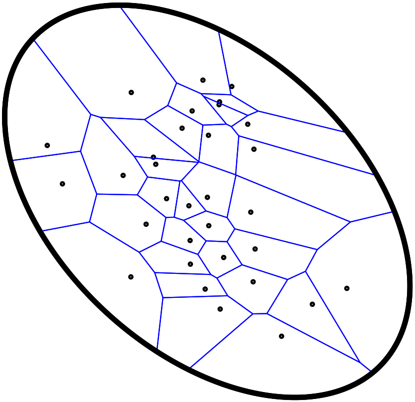
arccos and arccosh are monotonically increasing functions.

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



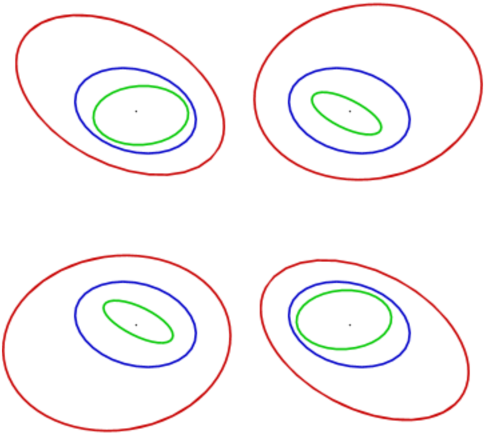
Hyperplanes (restricted to the domain)

Cayley-Klein Voronoi diagrams are affine



Can be computed from equivalent (clipped) power diagrams [2, 5]
https://www.youtube.com/watch?v=YHJLq3-RL58

Cayley-Klein balls



Blue: Mahalanobis Red: elliptic Green: Hyperbolic

Cayley-Klein balls have Mahalanobis ball shapes with displaced centers

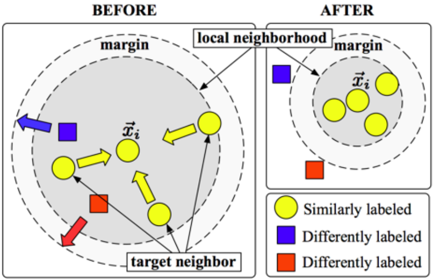
Learning curved Mahalanobis metrics

Large Margin Nearest Neighbors [8], LMNN

Learn Mahalanobis distance $M = L^\top L \succ 0$ 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*, $\epsilon_{\text{push}}(L)$
$$\mathcal{R} = \{(x_i, x_j, x_l) \ : \ (x_i, x_j) \in \mathcal{S} \text{ and } y_i \neq y_l\}$$



http://www.cs.cornell.edu/~kilian/code/lmnn/lmnn.html

LMNN: Cost function and optimization

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

$$
\begin{aligned}
\epsilon_{\text{pull}}(L) &= \Sigma_{i,i\to j}\|L(x_i - x_j)\|^2, \\
\epsilon_{\text{push}}(L) &= \Sigma_{i,i\to j}\Sigma_j(1-y_{il})\left[1 + \|L(x_i - x_j)\|^2 - \|L(x_i - x_l)\|^2\right]_+, \\
\epsilon(L) &= (1-\mu)\epsilon_{\text{pull}}(L) + \mu\epsilon_{\text{push}}(L)
\end{aligned}
$$

$i \to j$: 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.
μ set by cross-validation

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

$$
\frac{\partial\epsilon}{\partial L} = (1-\mu)\Sigma_{i,i\to j}C_{ij} + \mu\Sigma_{(i,j,l)\in\mathcal{R}_t}(C_{ij} - C_{il})
$$

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 $S = L^\top D L$ (with $L \succ 0$) and perform a gradient descent on L with the following gradient:

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

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 $\mathcal{P}$).

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

  with $\kappa > 1$.

- At iteration t, it may happen that $\mathcal{P} \notin \mathbb{D}_{S_t}$ since we do not know the optimal learning rate γ. When this happens, we reduce $\gamma \leftarrow \frac{\gamma}{2}$, otherwise we let $\gamma \leftarrow 1.01\gamma$.

Curved Mahalanobis learning: Results

Experimental results on some UCI data-sets

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

For classification, enough to consider $\kappa \in \{-1, 0, +1\}$

Spectral decomposition and fast proximity queries

- Avoid to compute d_E or d_H for arbitrary S
- Apply *spectral decomposition* (elliptic case $S = L^\top L$, or hyperbolic case $S = L^\top DL$) and perform coordinate changes so that we consider the canonical metric distances:

$$
\begin{aligned}
d_E(x', y') &= \arccos\left(\frac{\langle x', y'\rangle}{\|x'\|\|y'\|}\right), \\
d_H(x', y') &= \mathrm{arccosh}\left(\frac{1 - \langle x', y'\rangle}{\sqrt{1 - \langle x', x'\rangle}\sqrt{1 - \langle y', y'\rangle}}\right)
\end{aligned}
$$

- 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 α

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

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.
Discrete & Computational Geometry, 44(2):281–307, 2010.

C. Gunn.
Geometry, Kinematics, and Rigid Body Mechanics in Cayley-Klein Geometries.
PhD thesis, Technische UniversitÃďt Berlin, 2011.

Frank Nielsen, Boris Muzellec, and Richard Nock.
Classification with mixtures of curved Mahalanobis metrics.
In *IEEE International Conference on Image Processing (ICIP)*, pages 241–245, Sept 2016.

Frank Nielsen and Richard Nock.
Hyperbolic Voronoi diagrams made easy.
In *IEEE International Conference on Computational Science and Its Applications (ICCSA)*, pages 74–80, 2010.

Frank Nielsen, Paolo Piro, and Michel Barlaud.
Bregman vantage point trees for efficient nearest neighbor queries.
In *2009 IEEE International Conference on Multimedia and Expo*, pages 878–881. IEEE, 2009.

Jürgen Richter-Gebert.
Perspectives on Projective Geometry: A Guided Tour Through Real and Complex Geometry.
Springer Publishing Company, Incorporated, 1st edition, 2011.

Kilian Q. Weinberger, John Blitzer, and Lawrence K. Saul.
Distance metric learning for large margin nearest neighbor classification.
In *In NIPS*. MIT Press, 2006.

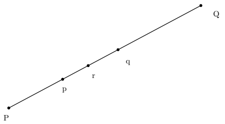
Eric P. Xing, Andrew Y. Ng, Michael I. Jordan, and Stuart Russell.
Distance metric learning, with application to clustering with side-information.
In *Advances in neural information processing systems*15*, pages 505–512. MIT Press, 2003.

Peter N. Yianilos.
Data structures and algorithms for nearest neighbor search in general metric spaces.
In *Proceedings of the Fourth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '93, pages 311–321, Philadelphia, PA, USA, 1993. Society for Industrial and Applied Mathematics.

Overview

Properties of the cross-ratio

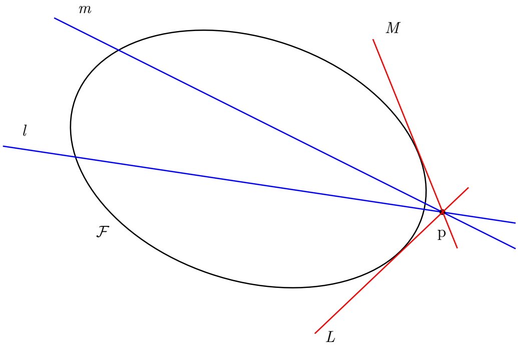
- $(p, p; P, P) = 1$
- $(p, q; Q, P) = \frac{1}{(p,q;P,Q)}$
- $(p, q; P, Q) = (p, r; P, Q) \cdot (r, q; P, Q)$ when r is collinear with p, q, P, Q



Measuring angles in Cayley-Klein geometries

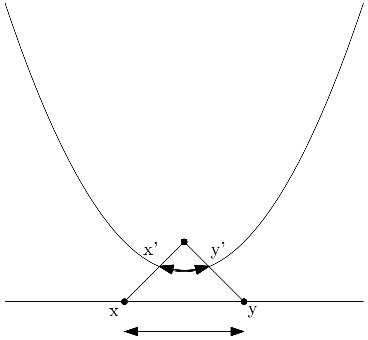
$$\mathrm{angle}(l, m) = c_{\mathrm{angle}} \log((l, m; L, M))$$

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

$$d_H(x, y) = \kappa \,\mathrm{arccosh}\left(\prec x', y' \succ\right)$$



Cayley-Klein Voronoi diagrams from (clipped) power diagrams

$$c_i = \frac{\Sigma p_i + a}{2\sqrt{S_{p_i p_i}}}$$

$$r_i^2 = \frac{\|\Sigma p_i + a\|^2}{4 S_{p_i p_i}} + \frac{a^\top p_i + b}{\sqrt{S_{p_i p_i}}}$$

Cayley-Klein balls have Mahalanobis ball shapes

Elliptic Cayley-Klein ball case:

$$\begin{aligned} \Sigma' &= \tilde{r}^2 \Sigma - a a^\top \\ c' &= \Sigma'^{-1}(b' a' - \tilde{r}^2 a) \\ r'^2 &= b'^2 - \tilde{r}^2 b + \langle c', c' \rangle_{\Sigma'} \end{aligned} \qquad \text{with} \qquad \begin{aligned} \tilde{r} &= \sqrt{S_{c,c}} \cos(r) \\ a' &= \Sigma c + a \\ b' &= a^\top c + b \end{aligned}$$

Cayley-Klein balls have Mahalanobis ball shapes

Hyperbolic Cayley-Klein ball case:

$$\begin{aligned} \Sigma' &= aa^\top - \tilde{r}^2\Sigma \\ c' &= \Sigma'^{-1}(\tilde{r}^2 a - b'a') \\ r'^2 &= \tilde{r}^2 b - b'^2 + \langle c', c' \rangle_{\Sigma'} \end{aligned} \quad \text{with} \quad \begin{aligned} \tilde{r} &= \sqrt{S_{c,c}}\cosh(r) \\ a' &= \Sigma c + a \\ b' &= a^\top c + b \end{aligned}$$

... 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: $S = O\Lambda O^\top$.

$$\Lambda = \mathrm{diag}(\Lambda_{1,1}, \ldots, \Lambda_{d+1,d+1})$$

- Canonical decomposition: $S = OD^{\frac{1}{2}} \begin{bmatrix} I & 0 \\ 0 & \lambda \end{bmatrix} D^{\frac{1}{2}} O^\top$, 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 $D_{d+1,d+1} = |\Lambda_{d+1,d+1}|$