UCL UNDERGRADUATE MATHEMATICS COLLOQUIUM

The Stiefel and Grassmann Manifolds

Orthonormal frames, their spans, and the map that forgets the basis

Yuvan Marimuthu · 28 October 2026

THE MAIN RESULT

Two orthonormal frames span the same subspace exactly when they differ by an orthogonal matrix.

$$\mathrm{St}(n, k) / \mathrm{O}(k) \cong \mathrm{Gr}(k, n)$$ Theorem 6.4

St(n, k): orthonormal frames · O(k): $k \times k$ orthogonal matrices · Gr(k, n): k-dimensional subspaces · the slash: frames differing by an orthogonal matrix count as one

Plan

1 A plane and its frames

2 Charts and manifolds

3 Constructing Stiefel

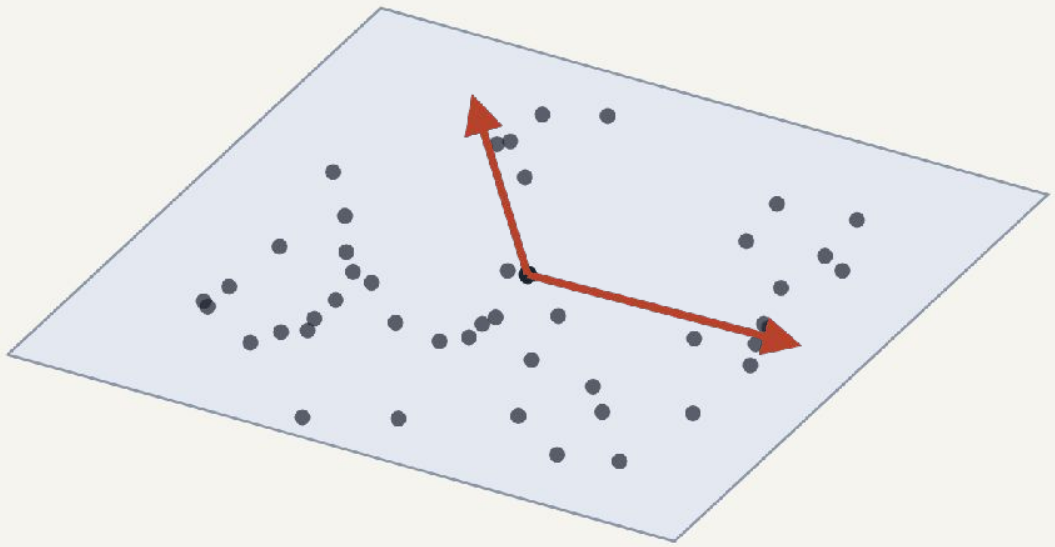
4 Coordinates for subspaces

5 The link and its proof

6 Application

Based on Segal’s notes, with optimisation from Boumal. Full proofs in the paper. No manifolds assumed.

A question from data



PRINCIPAL COMPONENT ANALYSIS

The plane through the origin minimising the sum of squared perpendicular distances.

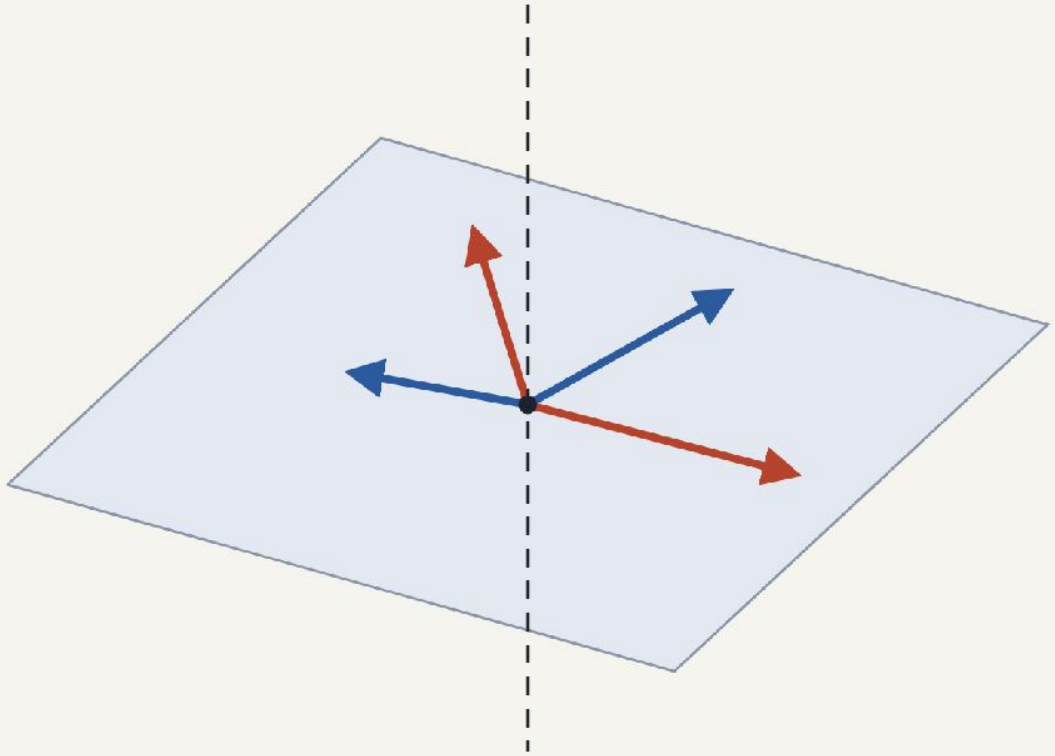
Represented by an orthonormal pair.

The representation is not unique.

1 · A plane and its frames

Two bases, one plane

Red and blue: two orthonormal bases of the same plane through the origin in $\mathbb{R}^3$



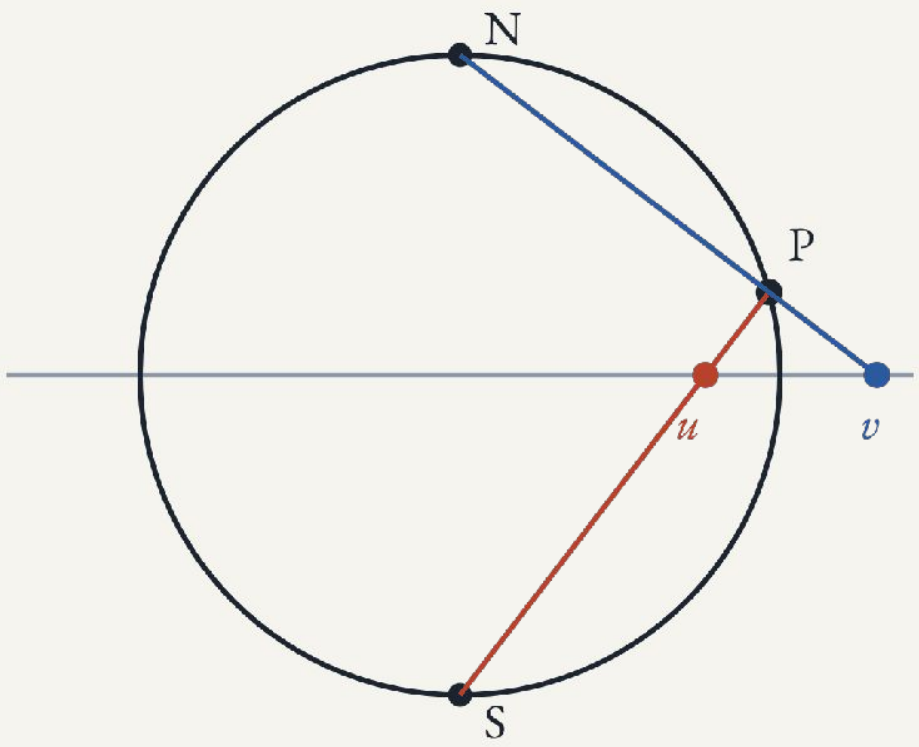
planes: 2

frames: 3

local dimensions

1 · A plane and its frames

Charts: flattening the circle



DEFINITION 2.1 · CO-ORDINATE CHART

A homeomorphism from an open subset of the space onto an open subset of $\mathbb{R}^d$.

The coordinate map and its inverse are both continuous.

From the south pole: $u = x / (1 + y)$

From the north pole: $v = x / (1 - y)$

$$u \cdot v = x^2 / (1 - y^2) = 1$$

The transition function

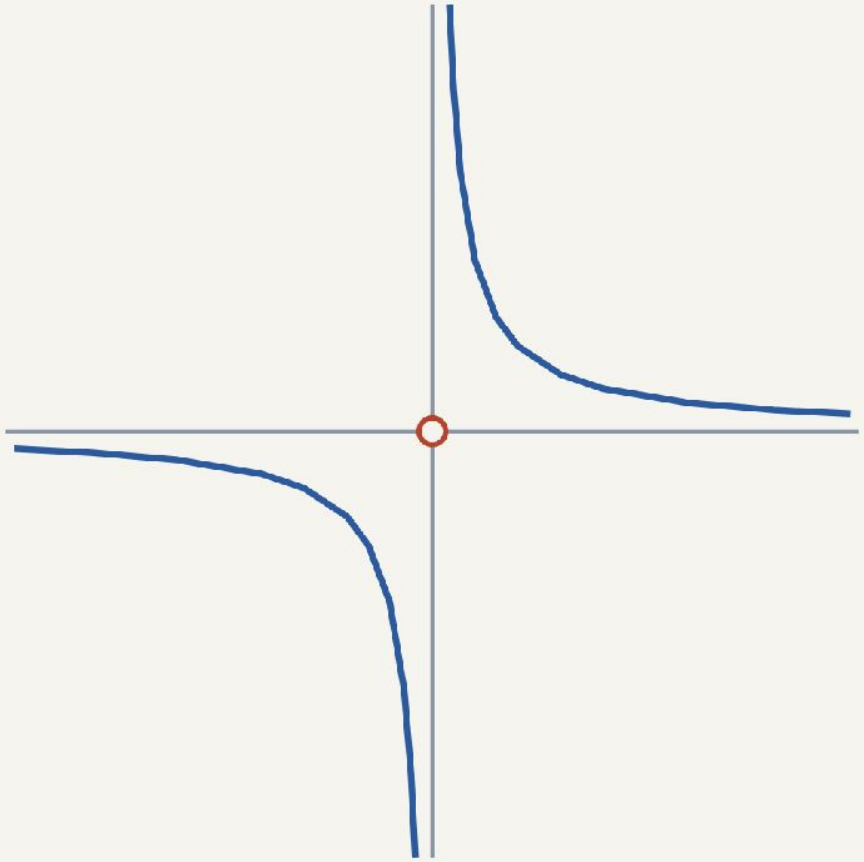
DEFINITION 2.3 · TRANSITION FUNCTION

$\varphi_2 \circ \varphi_1^{-1}$ on $\varphi_1(U \cap V)$

One chart's coordinates written in terms of the other's.

$$u \mapsto 1/u$$

Translates one chart's coordinate into the other's where both apply. Smooth wherever $u \neq 0$.



Manifolds

TOPOLOGICAL MANIFOLD · DEFINITION 2.2

A d-dimensional manifold admits local coordinates in $\mathbb{R}^d$ near every point, and is Hausdorff and second countable.

Hausdorff: distinct points have disjoint open neighbourhoods.

Second countable: the topology has a countable basis.

Both hold for subsets of Euclidean space.

SMOOTH MANIFOLD · DEFINITIONS 2.4 AND 2.7

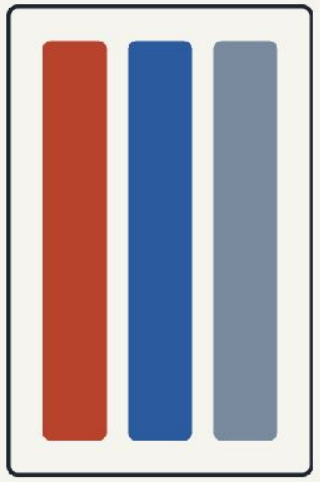
Atlas: charts covering the space. Smooth: every transition is infinitely differentiable.

Circle: $u \mapsto 1/u$ is smooth.

Changing chart does not change whether a function is smooth.

The Stiefel manifold

$\mathrm{St}(n, k) = \{ X \in \mathbb{R}^{n \times k} : X^{\mathsf{T}} X = I_k \}$. Columns = the frame.



X: $n \times k$, columns = the frame

$$X^{\mathsf{T}} X = I_k$$

1	0	0
0	1	0
0	0	1

Entry (i, j) = column i · column j

$k = 1$: the sphere S^{n-1}

$k = n$: the group $\mathrm{O}(n)$

Compact: closed, and every X has squared Frobenius norm k, so Heine–Borel applies

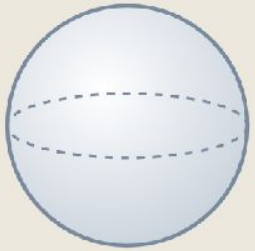
Why is it a manifold?

REGULAR LEVEL SETS · THEOREM 2.16

$F : \mathbb{R}^a \to \mathbb{R}^b$ smooth (finite-dimensional spaces), with $DF(x)$ onto at every solution of $F(x) = c$.

If the solution set is nonempty, it is a manifold of dimension $a - b$.

Tangent space at x = ker $DF(x)$. Without the onto condition: $\|x\|^2 = 0$ is a single point, and $DF = 0$ there.



EXAMPLE 2.24 · THE SPHERE

$$F(x) = \|x\|^2,\ DF(x)v = 2x \cdot v, \text{ onto at unit } x$$

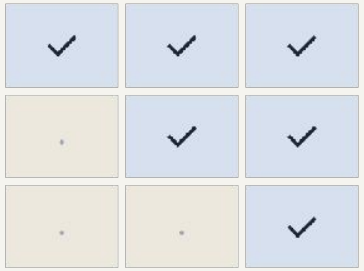
$$\dim S^2 = 3 - 1 = 2$$

How big is $\mathrm{St}(n, k)$?

$$h(X) = X^\mathsf{T}X - I_\mathbf{k}$$

nk entries. How many equations?

$X^\mathsf{T}X$ is symmetric: target $\mathrm{Sym}(k)$, of dimension $k(k+1)/2$.



THE DERIVATIVE IS ONTO

$$Dh(X)[V] = X^\mathsf{T}V + V^\mathsf{T}X$$

For symmetric S, take $V = XS/2$: then $Dh(X)[V] = S$.

THEOREM 4.3

$$\dim \mathrm{St}(n, k) = nk - k(k+1)/2$$

Two-vector frames in $\mathbb{R}^3$: $6 - 3 = 3$

The Grassmannian

DEFINITION 5.1

$$\mathrm{Gr}(k, n) = \{\, k\text{-dimensional subspaces of } \mathbb{R}^n \,\}$$

So far, a set

No topology yet, and no smooth structure.

Coordinates first

PSEUDO-CHARTS · SEGAL, DEFINITION 2.10

Bijections from subsets of the set onto open subsets of Euclidean space, before any topology is chosen.

They must cover the set.

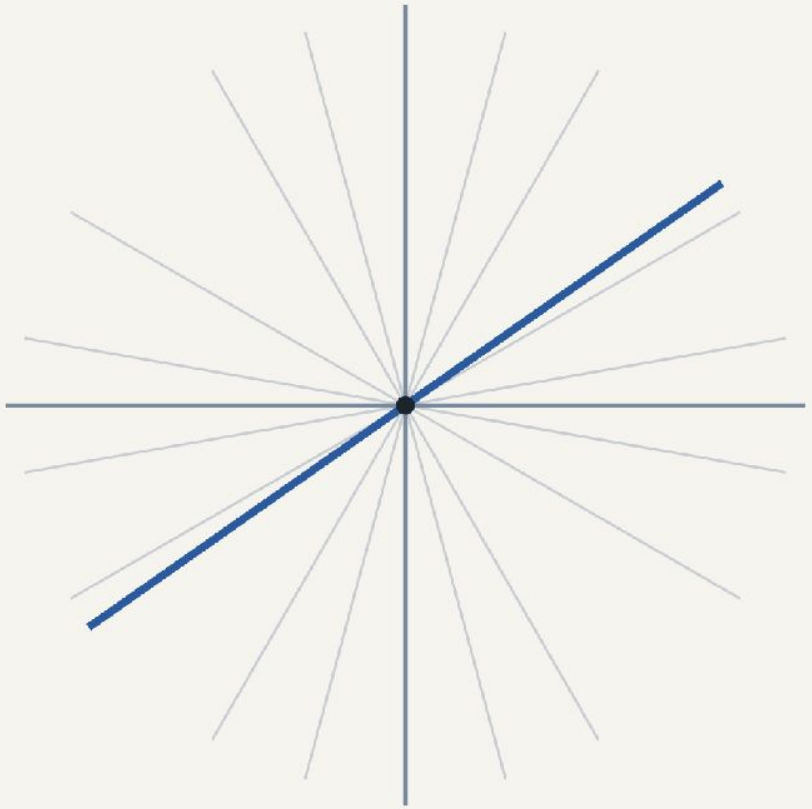
PROPOSITION 2.11

Open overlap images and continuous transitions ⟹ a unique topology in which they become charts.

COROLLARY 2.12

Smooth transitions, Hausdorff and second countable ⟹ a smooth manifold.

Lines through the origin in $\mathbb{R}^2$



$y = bx$

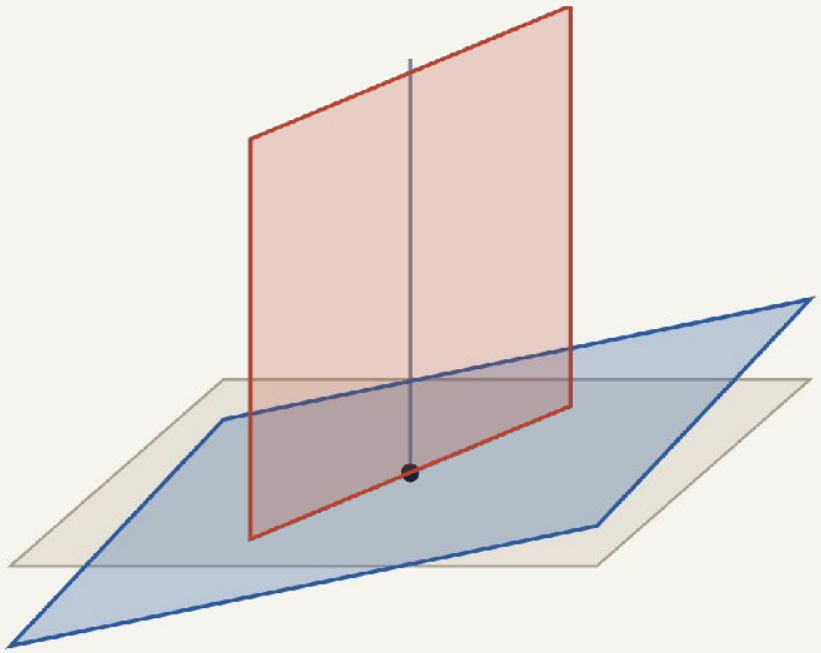
Every line except the vertical one has a slope b.

$x = cy$

Every line except the horizontal one has a number c.

$c = 1/b$ circle: $v = 1/u$

A plane as a graph



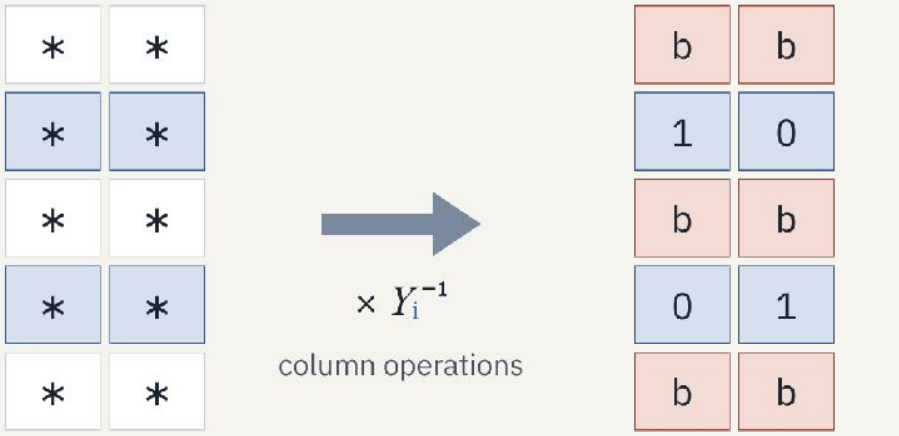
$$z = ax + by$$

$$Y = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ a & b \end{bmatrix}$$

The columns span the vectors $(x, y, ax + by)$. The pair (a, b) is the coordinate.

Planes containing the z-axis: choose a different pair of independent variables.

Coordinates for any subspace



Y, rows I in blue

rows $I \to$ identity, rest $\to B$

COORDINATES FOR A SUBSPACE · DEFINITION 5.3

Basis matrix Y. Choose k rows forming an invertible block Y_i, and multiply on the right by Y_i^{-1}. Same span; the other rows give B.

First k rows: $(I_k ; B)$ spans the vectors $(u ; Bu)$, the graph of $u \mapsto Bu$

Well defined: $YG(Y_iG)^{-1} = YY_i^{-1}$

Covering: a rank-k matrix has an invertible $k \times k$ row block

Every B arises: $(I_k ; B)$ is a basis, so each chart is onto all $(n - k) \times k$ matrices, an open set

$$\dim = k(n - k)$$

Overlaps and the two conditions

TRANSITIONS · LEMMA 5.5

$$(I_2\,;\,B)\,B^{-1} = (B^{-1}\,;\,I_2)$$

$B \mapsto B^{-1}$, as $u \mapsto 1/u$

In general: multiplication and inversion, on domains where a determinant is nonzero. Open domains, smooth transitions.

HAUSDORFF · PROPOSITION 5.7

$$P(S) = XX^{\mathrm{T}}$$

Continuous; distinct subspaces give distinct projectors. Separate them, then take preimages.

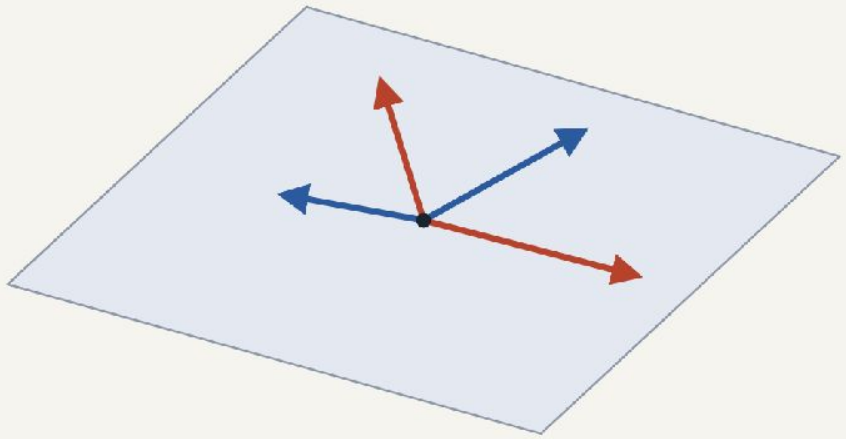
Second countable: finitely many charts.

THEOREM 5.8

$$\dim \mathrm{Gr}(k, n) = k(n-k)$$

Planes in $\mathbb{R}^3$: 2

Forgetting the basis



DEFINITION 6.1 · COLUMN-SPACE MAP

$$\pi : \mathrm{St}(n, k) \to \mathrm{Gr}(k, n)$$

$$\pi(X) = \mathrm{col}(X)$$

Onto: every subspace has an orthonormal basis. Continuous: in the chart for rows I, col(X) has coordinate (other rows of X) · X_I^{-1}, built from inversion, on the open set where X_I is invertible. Which frames have the same span?

Which frames have the same span?

PROPOSITION 6.2

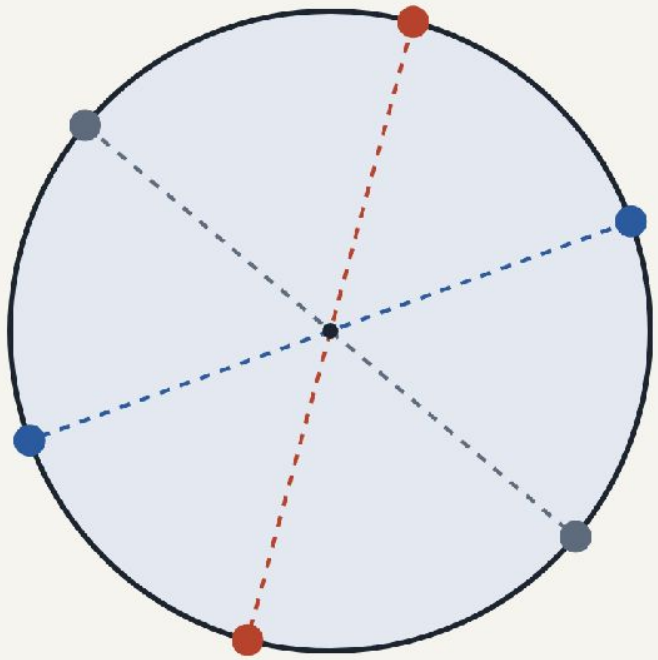
$\pi(X) = \pi(X')$ if and only if $X' = XQ$ for some $Q \in \mathrm{O}(k)$.

1. Same column space, so $X' = XQ$ with Q invertible.

2. $I = X'^{\mathsf{T}}X' = Q^{\mathsf{T}}X^{\mathsf{T}}XQ = Q^{\mathsf{T}}Q$, so $Q \in \mathrm{O}(k)$.

Conversely, $(XQ)^{\mathsf{T}}(XQ) = I_k$ with the same span. The orbit of X: $\{\, XQ : Q \in \mathrm{O}(k) \,\}$.

The case $k = 1$



$\mathrm{St}(n, 1) = S^{n-1}$, the sphere

$\mathrm{O}(1) = \{ +1, -1 \}$

Orbit of a unit vector: $\{ x, -x \}$

$\mathrm{St}(n, 1)/\mathrm{O}(1) = \mathbb{R}\mathrm{P}^{n-1}$

The quotient

$\mathrm{St}(n, k)/\mathrm{O}(k)$ is the set of orbits. A collection of orbits is open when all their frames form an open subset of Stiefel.

$$[X] \mapsto \mathrm{col}(X) \text{ is a bijection onto } \mathrm{Gr}(k, n).$$

LEMMA

A continuous bijection from a compact space to a Hausdorff space is a homeomorphism.

Closed ⊂ compact is compact → its image is compact → compact ⊂ Hausdorff is closed. So the inverse is continuous.

Continuous: π is, and the quotient topology. Compact: the image of compact Stiefel. Hausdorff: Gr, by projection matrices.

$$\mathrm{St}(n, k)/\mathrm{O}(k) \cong \mathrm{Gr}(k, n) \cdot \text{Theorem 6.4}$$

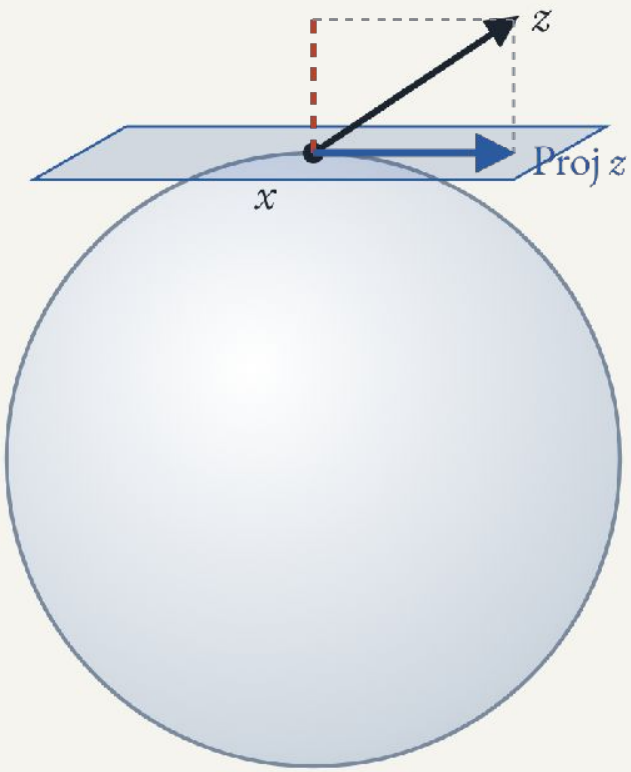
Application: orthogonal recurrent networks

$$h_t = \sigma(W h_{t-1} + U x_t + b)$$

$W^{\mathsf{T}}W = I_n$ puts W in $\mathrm{O}(n) = \mathrm{St}(n, n)$. The linear map $h \mapsto Wh$ preserves lengths; the full update need not.

Training changes W to reduce a loss. An arbitrary matrix update usually leaves $\mathrm{O}(n)$.

Staying on the manifold



TANGENT SPACE · PROPOSITION 4.5

Velocities of smooth curves that stay on the manifold.

$$W^{\mathrm{T}}V + V^{\mathrm{T}}W = 0$$

The kernel of the derivative, as in the regular level set theorem.

PROJECTED GRADIENT · FROBENIUS INNER PRODUCT · COROLLARY 4.7

$$\operatorname{grad} L(W) = G - W\operatorname{sym}(W^{\mathrm{T}}G)$$

$\operatorname{sym}(A) = (A + A^{\mathrm{T}})/2$. On the sphere: remove the radial part.

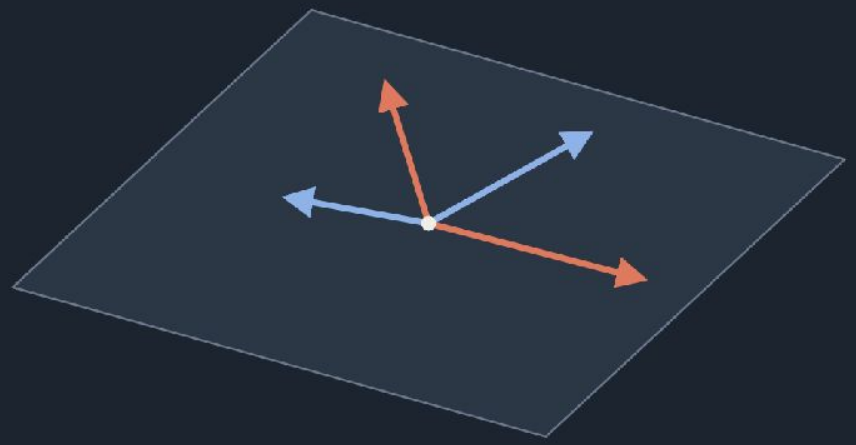
PROVED IN THE PAPER

The polar retraction turns a tangent step into a point on the manifold (Proposition 4.9).

PLANNED

The implementation and the experiments.

The plane from the beginning



Red and blue: two points of St(3, 2).

Their common span: one point of Gr(2, 3).

What we built

FRAMES

$X^{\mathrm{T}}X = I_{\mathbf{k}}$

SUBSPACES

Locally, graphs of linear maps $u \mapsto Bu$

COLUMN-SPACE MAP

Fibres: the orthogonal changes of basis $\{ XQ \}$

$$\mathrm{St}(n, k)/\mathrm{O}(k) \cong \mathrm{Gr}(k, n)$$

Questions

Stiefel and Grassmann Manifolds: A Geometric Foundation for Optimisation with Orthogonality Constraints

Full proofs in the paper: [link to paper]

Yuvan Marimuthu · UCL

With thanks to Professor Ed Segal.