Conditions for the Well-Posedness of Non-Negative Matrix Factorization

ICA Workshop - NIPS 2003

David Donoho, Victoria Stodden

mailto: vcs@stat.stanford.edu

December 13, 2003

Lee & Seung 1999

letters to nature

larvae collected randomly in the field (2° 48.12' N, 41° 40.33' E) by SCUBA. Between 5 and 10 juveniles were recruited successfully in each of 15, 1 l polystyrene containers ($n = 15$), the bottom of which was covered with an acetate sheet that served as substratum for sponge attachment. Containers were then randomly distributed in 3 groups, and sponges in each group were reared for 14 weeks in 3 different concentrations of $Si(OH)_4$: 0.741 ± 0.133, 30.235 ± 0.287 and 100.041 ± 0.760 μM (mean ± s.e.). All cultures were prepared using 0.22 μm polycarbonate-filtered seawater, which was collected from the sponge habitat, handled according to standard methods to prevent Si contamination[29] and enriched in dissolved silica, when treatments required, by using Na_2SiF_6. During the experiment, all sponges were fed by weekly addition of 2 ml of a bacterial culture (40–60 × 10^6 bacteria ml^{-1}) to each container[30]. The seawater was replaced weekly, with regeneration of initial food and $Si(OH)_4$ levels. The concentration of $Si(OH)_4$ in cultures was determined on 3 replicates of 1 ml seawater samples per container by using a Bran-Luebbe TRAACS 2000 nutrient autoanalyser. After week 5, the accidental contamination of some culture containers by diatoms rendered subsequent estimates of Si uptake by sponges unreliable, so we discarded them for the study.

For the study of the skeleton, sponges were treated according to standard methods[30] and examined in a Hitachi S-2300 scanning electron microscope (SEM).

Received 21 April; accepted 16 August 1999.

1. Hartman, W. D., Wendt, J. W. & Wiedenmayer, F. Living and fossil sponges. Notes for a short course. *Sedimenta* 8, 1–274 (1980).
2. Ghiold, J. The sponges that spanned Europe. *New Scient.* 129, 58–62 (1991).
3. Leinfelder, R. R. Upper Jurassic reef types and controlling factors. *Profil* 5, 1–45 (1993).
4. Wiedenmayer, F. Contributions to the knowledge of post-Paleozoic neritic and archibental sponges (Porifera). *Schweiz. Paläont. Abh.* 116, 1–147 (1994).
5. Lévi, C. in *Fossil and Recent Sponges* (eds Reitner, J. & Keupp, H.) 72–82 (Springer, New York, 1991).
6. Moret, L. Contribution à l'étude des spongiaires siliceux du Miocene de l'Algerie. *Mém. Soc. Géol. Fr.* 1, 1-27 (1924).

Learning the parts of objects by non-negative matrix factorization

Daniel D. Lee* & H. Sebastian Seung*†

* *Bell Laboratories, Lucent Technologies, Murray Hill, New Jersey 07974, USA*
† *Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

Is perception of the whole based on perception of its parts? There is psychological[1] and physiological[2,3] evidence for parts-based representations in the brain, and certain computational theories of object recognition rely on such representations[4,5]. But little is known about how brains or computers might learn the parts of objects. Here we demonstrate an algorithm for non-negative matrix factorization that is able to learn parts of faces and semantic features of text. This is in contrast to other methods, such as principal components analysis and vector quantization, that learn holistic, not parts-based, representations. Non-negative matrix factorization is distinguished from the other methods by its use of non-negativity constraints. These constraints lead to a

Claim

Their Paper:

- Vague relation between NMF and "Decomposing into Parts."
- NMF a vague idea.

Our talk:

- Gives a theoretical framework that validates NMF in specific settings.
- Identifies under what conditions NMF works.

The Problem:

$$X = A\Psi$$

X is $n \times p$ $\quad$ A is $n \times r$

Ψ is $r \times p$

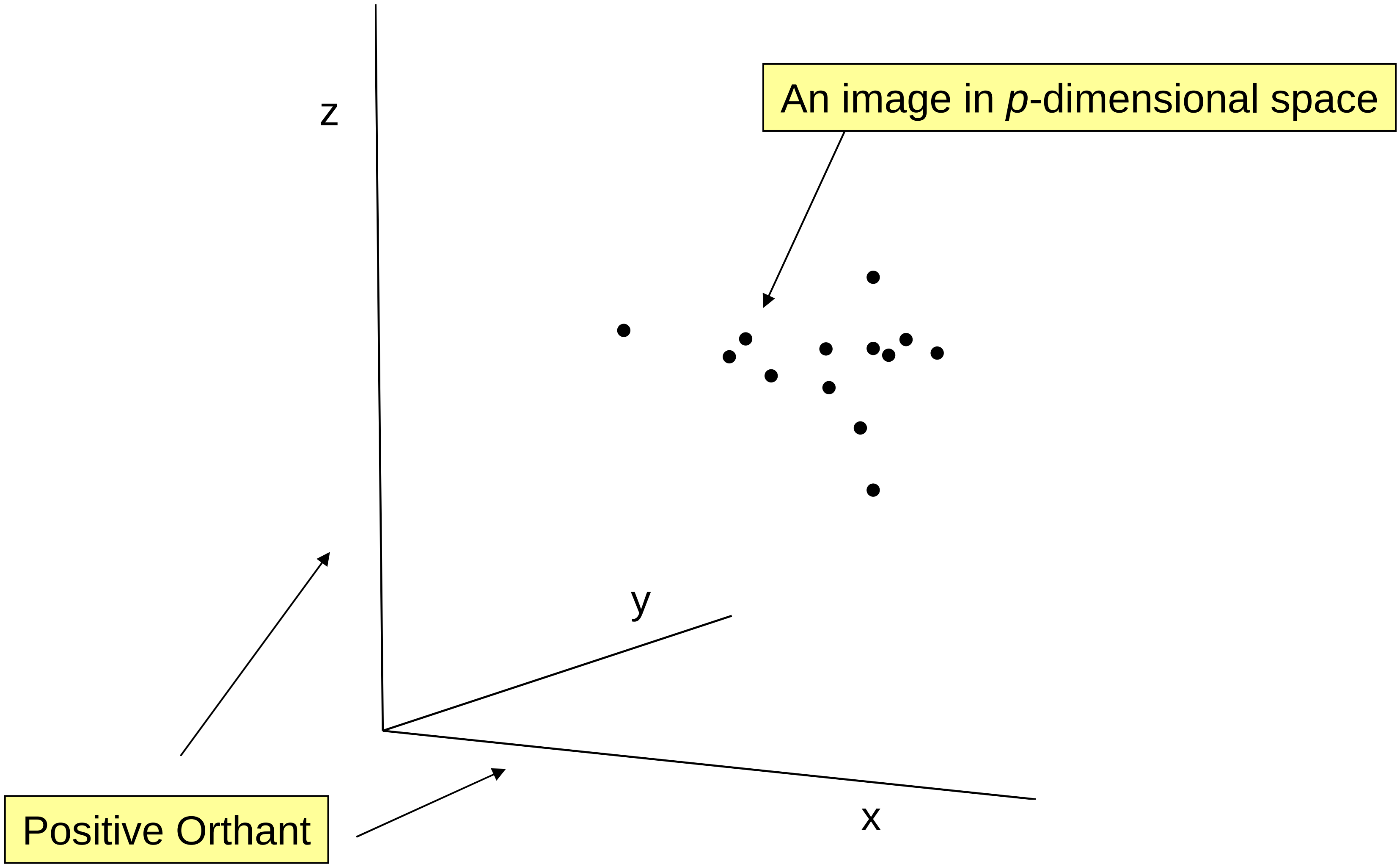
$$A \geq 0, \Psi \geq 0$$

Does this make sense?

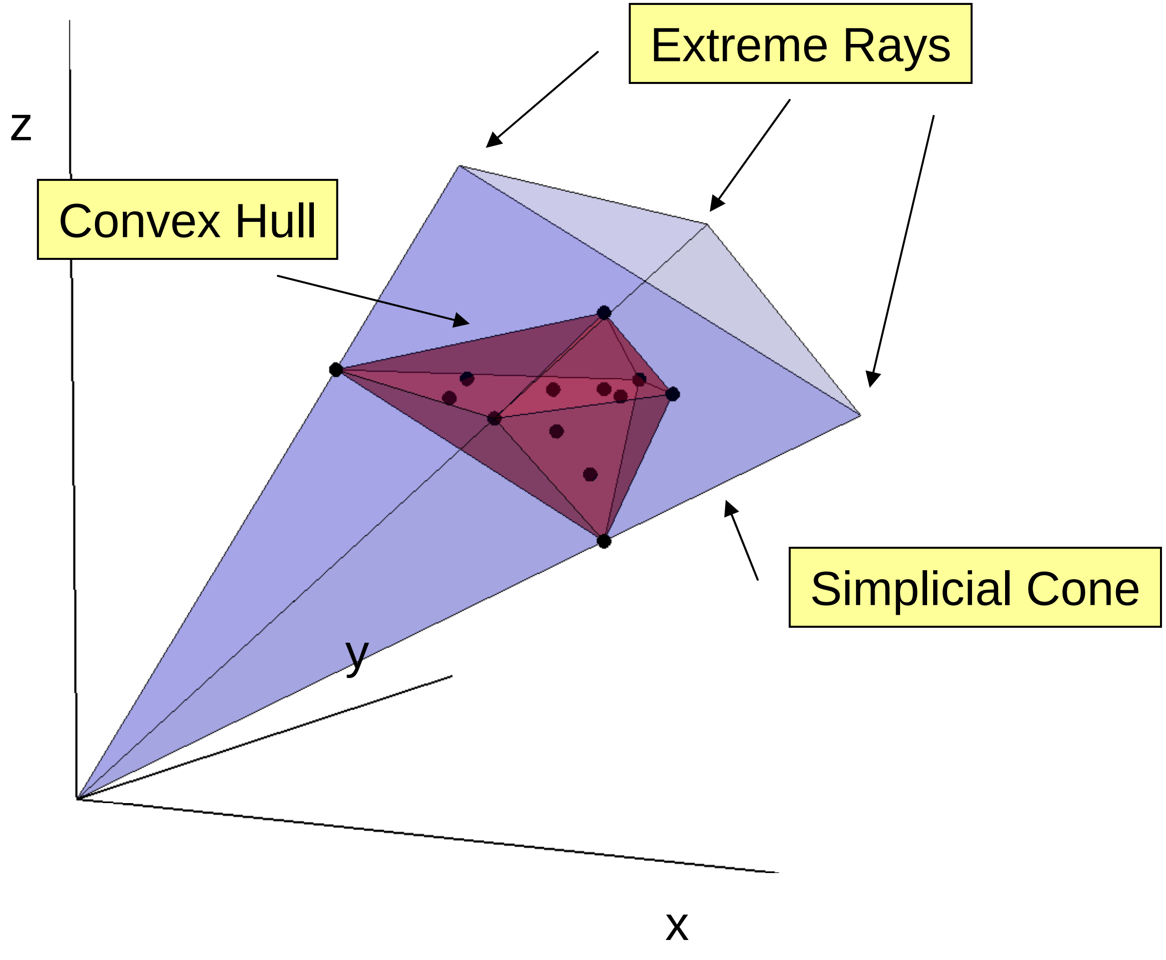
- **When is the factorization *unique*?**
- **When does NMF find the *correct* decomposition?**

- Plumbley (2001) proposed an ICA model that solves the NMF problem, where $n \gg p = r$, the A_{ij} are *independent* random variables, and the basis functions are non-negative and linearly independent.
- Parts-articulation models behind the Lee & Seung proposal do not have independent components.

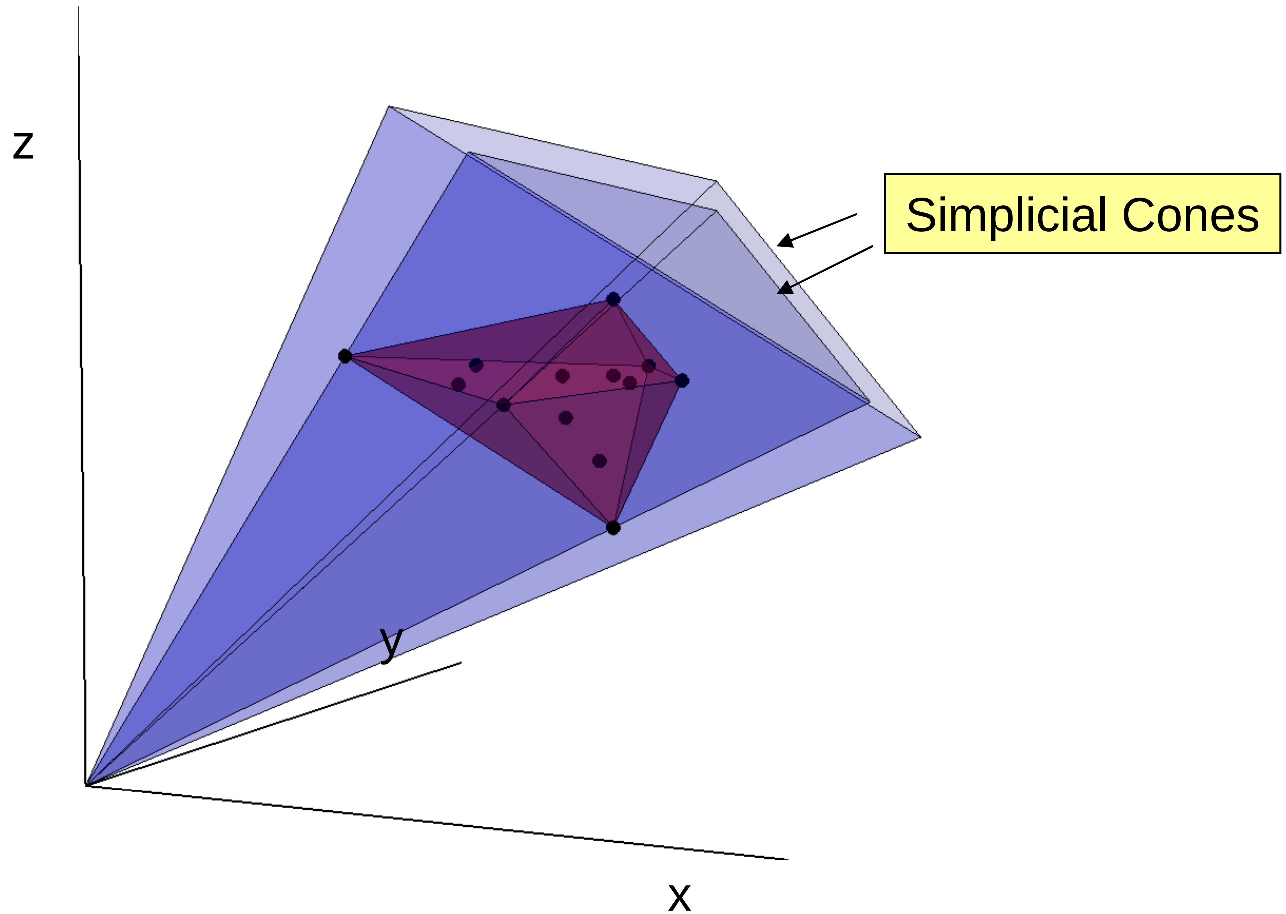
Geometric Reinterpretation



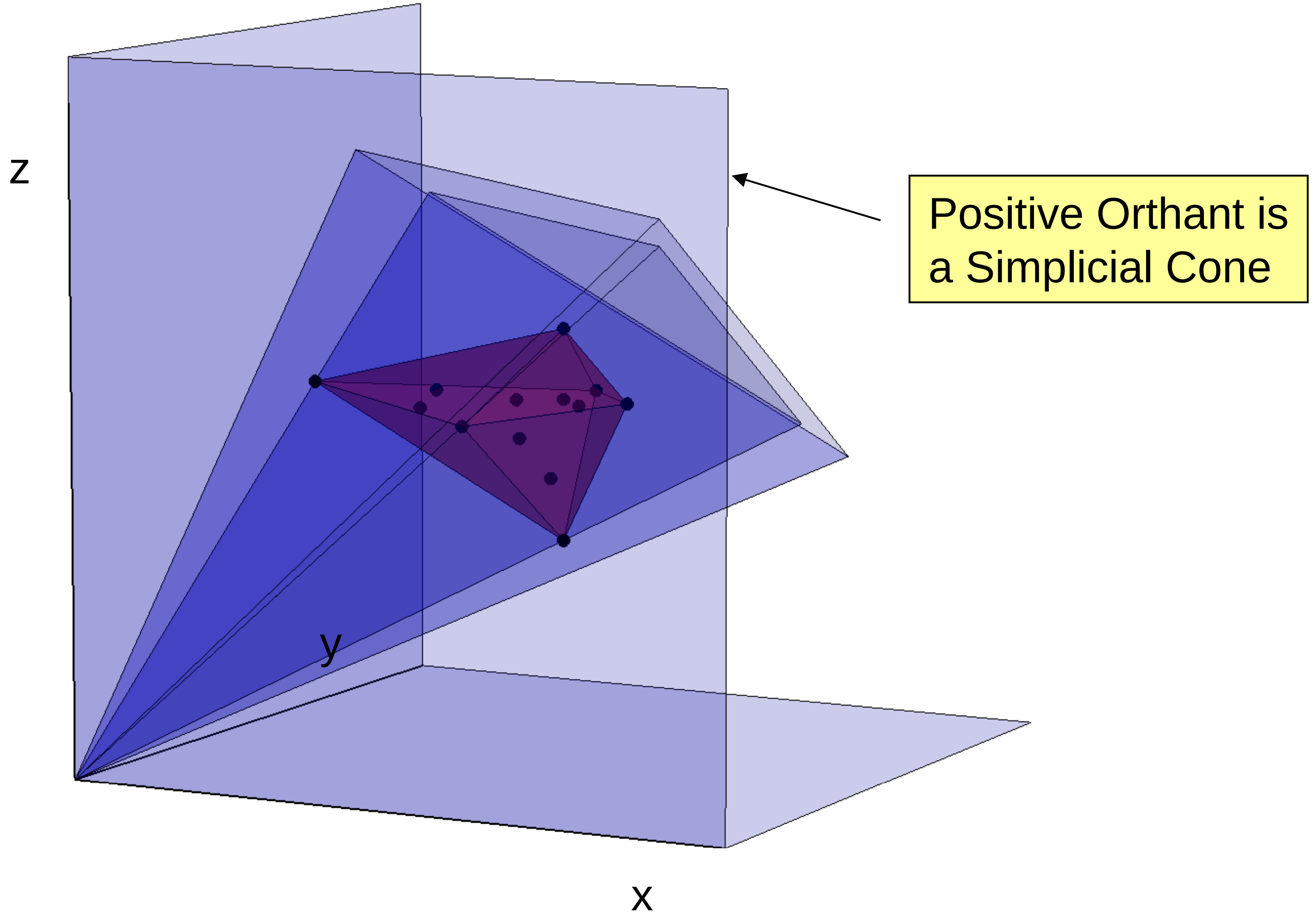
Geometric Interpretation of NMF



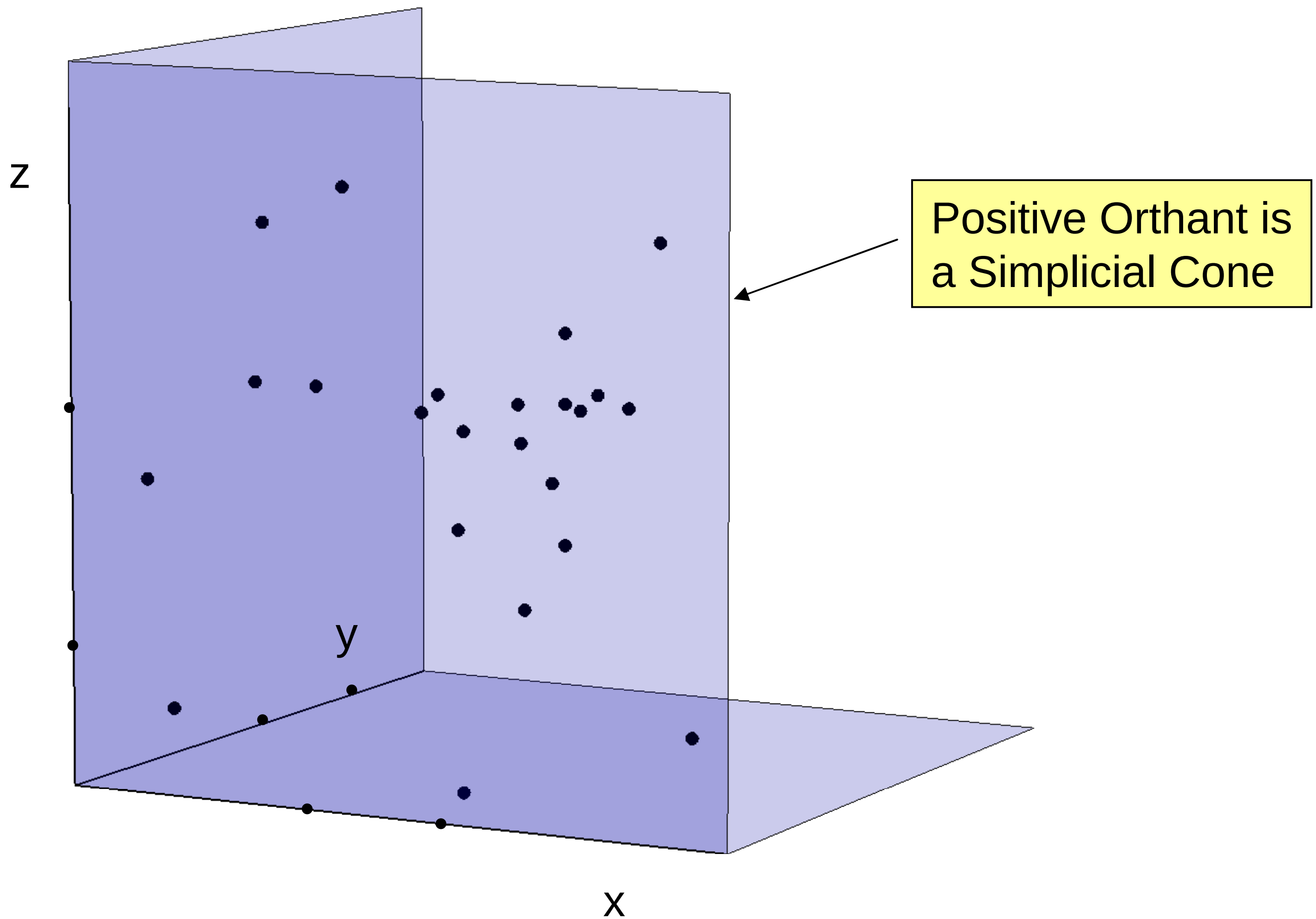
Not Unique!



Not Unique!

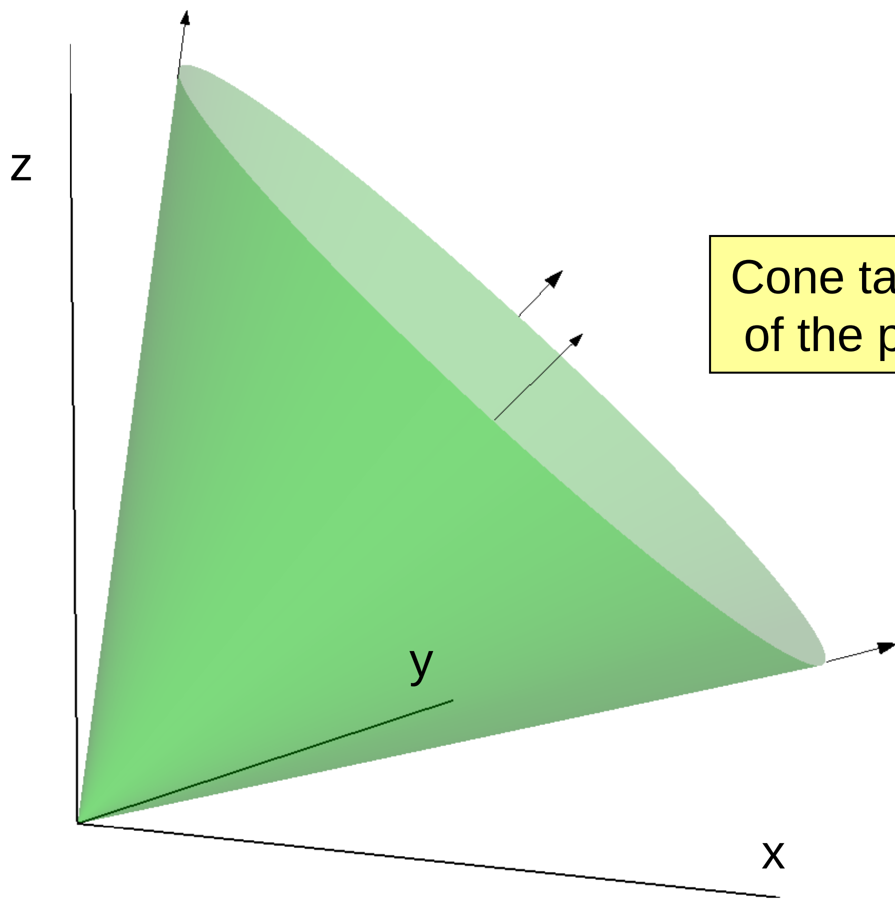


Example of Uniqueness I



Example of Uniqueness II

$$C = \{x : x'\mathbf{1} \geq \sqrt{p-1} \, \| x \| \}$$



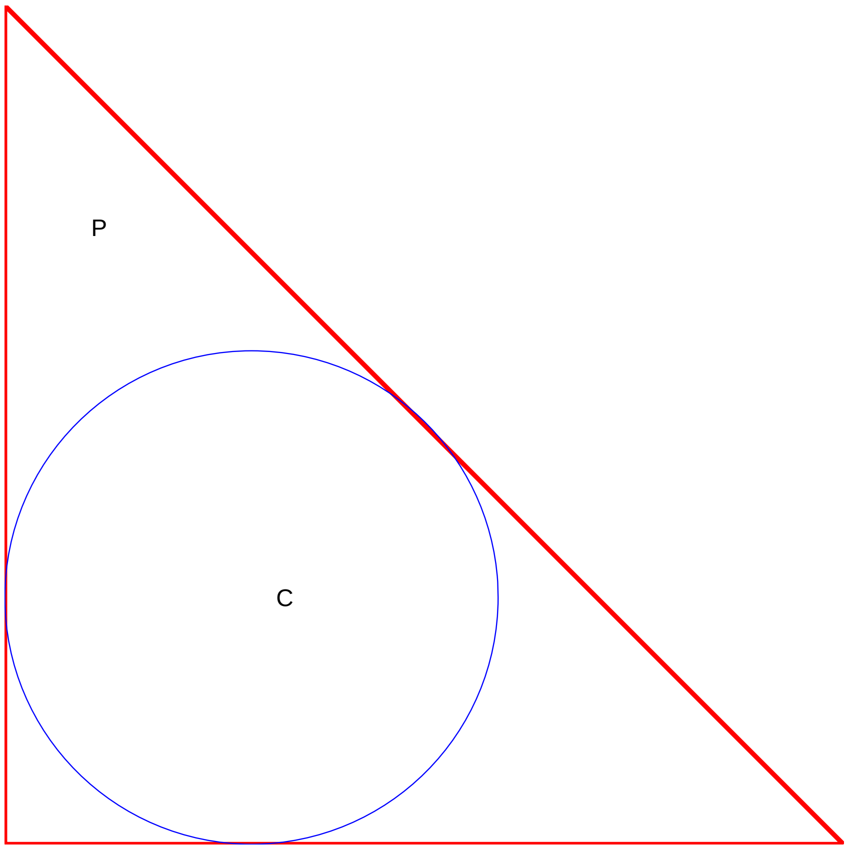
Cone tangent to the planes of the positive orthant.

Convex Duality

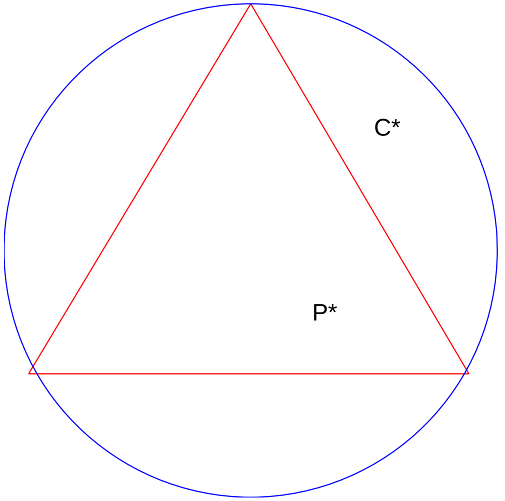
- Notion of primal space and dual space.
- Points in one space map to hyperplanes in the other.
- Convex cones map to convex cones.
- Inclusions reverse under duality.

Reinterpretation using Duality

Primal Space



Dual Space



Theorem

Given a Separable Factorial Articulation Database, there is a unique simplicial hull with $r=AP$ generators which contains all the data points and is contained in the intersection of the positive orthant and the linear span of the generators.

This unique NMF solution is visibly a correct decomposition into parts.

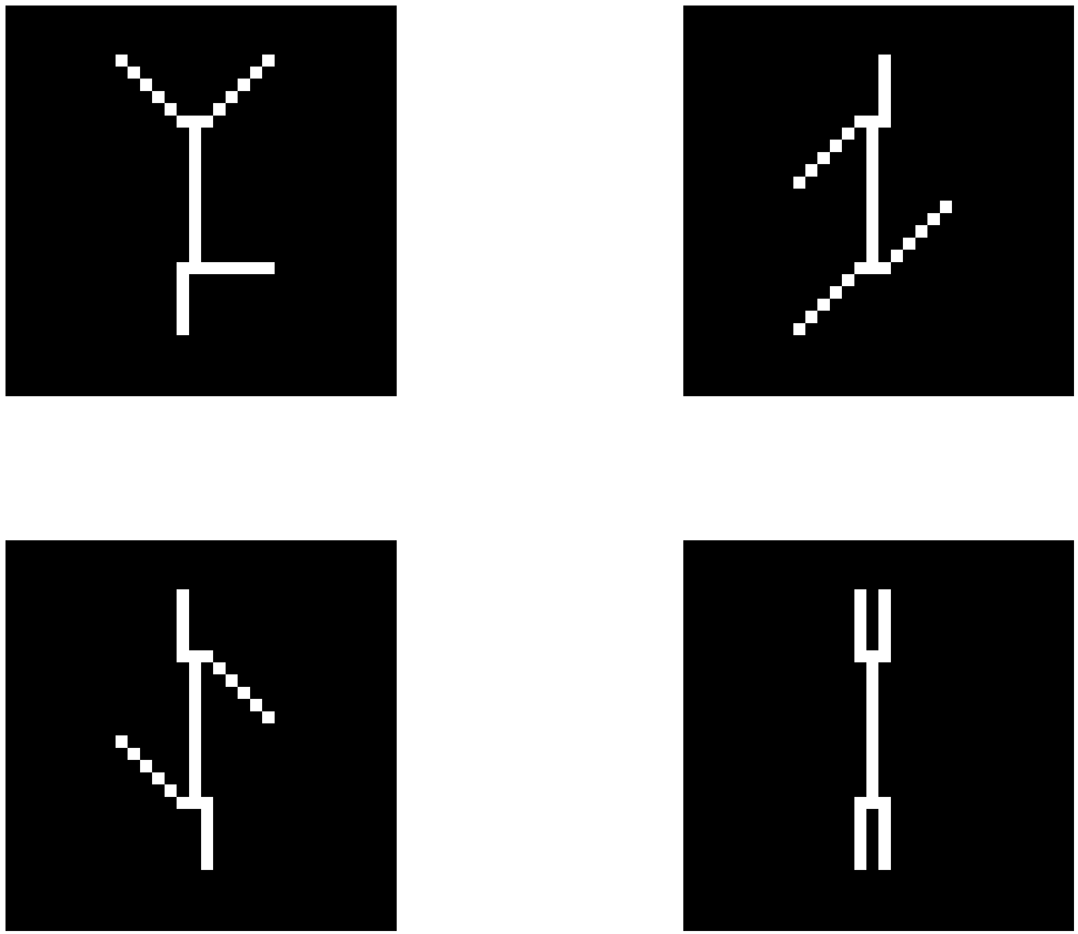
Separable Factorial Articulation Families

Definition:

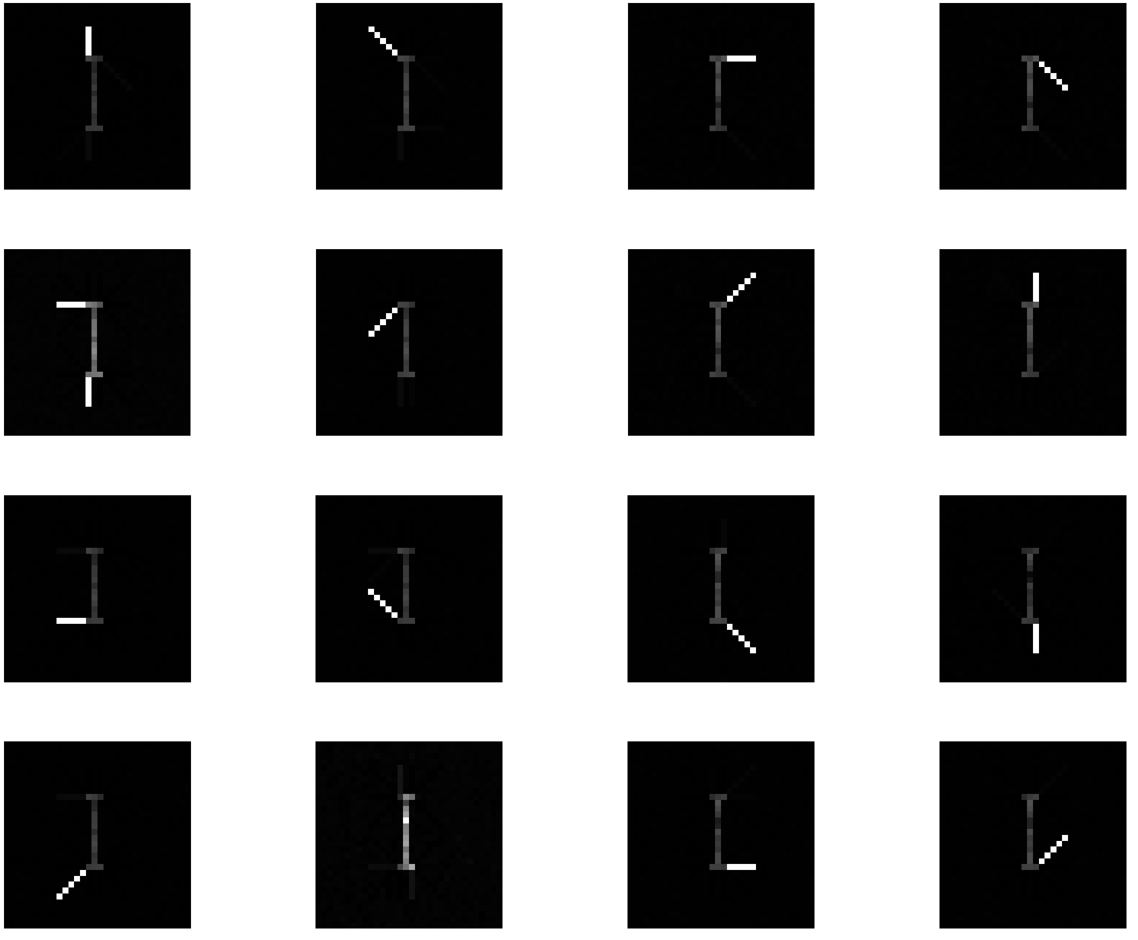
- A structured family of images, *Xi*, where *P* is the number of parts and *A* is the number of articulations.
- Each image, X_i, equals a superposition of every part in only one of its articulations.
- Existence of tell-tale pixel for each part/articulation.
- Every possible combination of every part and every articulation occurs.

Consequence: each part can only appear in one articulation at a time, implying negative dependence.

Example: `Swimmer`



NMF Parts Recovery



Conclusions

- The geometric viewpoint explains the non-well-posedness of NMF in general, and the well-posedness in a specific setting.
- NMF makes sense, in theory, for Separable Factorial Articulation Families.
- The theory correctly predicts the success of non-negative matrix factorization.