Solving the 2-norm k-**Hyperplane Clustering Problem** via Multi-Norm Formulations

Stefano Coniglio

Department of Economics
University of Bergamo
Bergamo, Italy

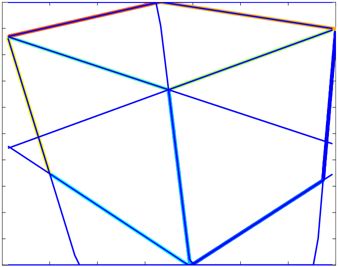
stefano.coniglio@unibg.it

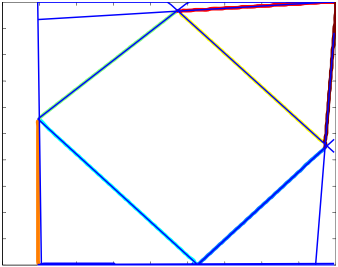
Hyperplane Clustering

In **classical clustering** (k-means), we cluster over **centroids**—0-dimensional subspaces of the same dimensionality as the data points.

In many applications, though, one seeks **relationship of *collinearity* or *coplanarity*** among the datapoints.

Example: recover the equations of the lines that make up a digital image.





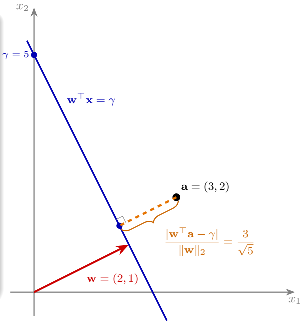
k-HC is the problem of **clustering points along hyperplanes** (n-dimensional subspaces):

Problem definition

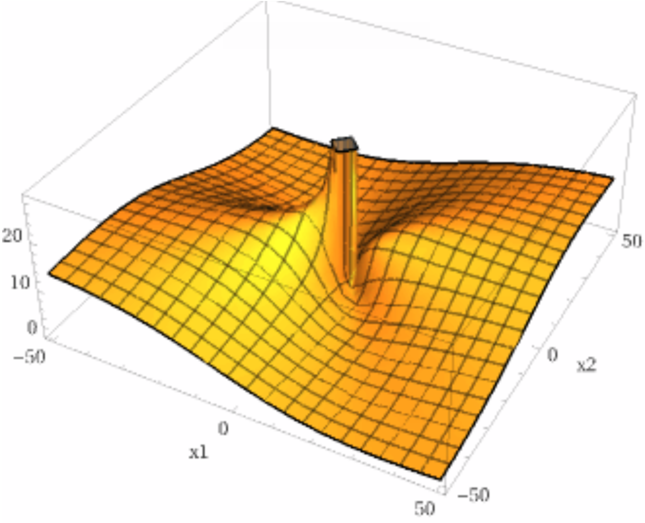
Given m points $\{a_1, \ldots, a_m\}$ in $\mathbb{R}^n$:

- determine the affine equations $w_j^\top x = \gamma_j$ of k **hyperplanes**, $j \in [k]$
- assign each point to one hyperplane

minimizing the sum of squared **2-norm orthogonal distances** $\left(\frac{|w_j^\top a_i - \gamma_j|}{\|w_j\|_2}\right)^2$



The point-to-hyperplane Euclidean distance entails a normalization by $\|w\|_2$ which makes it **highly nonconvex**—k-HC is **very** hard to solve to global optimality.

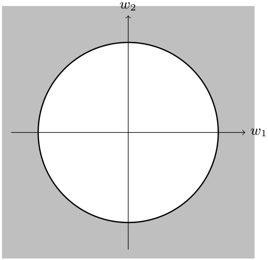


WolframAlpha:

`plot3d (3x1 + 2x2 - 5)^2 / (x1^2 + x2^2) for x1 from -50 to 50, x2 from -50`

The challenge: the 2-norm constraint(s)

W.l.o.g., we can move the scaling away from the objective function into the $\|w_j\|_2 = 1$ which, in turn, can be relaxed w.l.o.g. to $\|w_j\|_2 \geq 1$.
With this, we obtain, for each normal vector w_j, a feasible region coinciding with the *complement* of the unit Euclidean ball: a **concave set**.



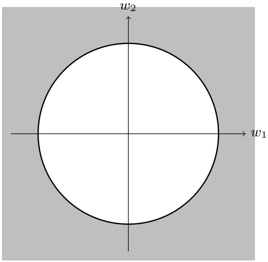
Whe solve the problem via **Spatial Branch-and-Bound** (SBB), which is based on building convex relaxations of the problem.

Proposition

Under midpoint branching, a nonzero global lower bound for k-HC requires is obtained after at least $2^{k(n-1)}$ branching nodes (no pruning takes place until then).

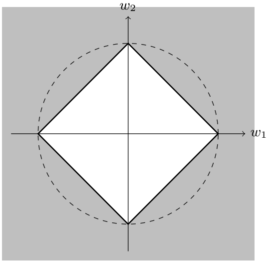
Research question

If we replace $\|w_j\|_2 \geq 1$ with *polyhedral* norm constraints (which are easy to express via 0–1 programming techniques), do we get a good approximation?



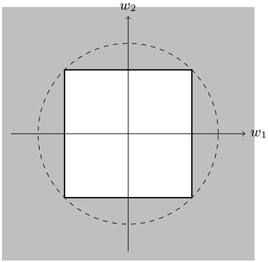
$$\|w\|_2 \geq 1$$

sphere
k-HC$_{(2,1)}$



$$\|w\|_1 \geq 1$$

(hyper)octahedron
k-HC$_{(\infty,1)})$



$$\|w\|_\infty \geq \frac{1}{\sqrt{n}}$$

hypercube
k-HC$_{(1,\frac{1}{\sqrt{n}})}$.

Norm equivalence ⇒ approximation guarantee

Theorem (approximation factors)

$$\frac{1}{n}\,\mathrm{OPT}(k\text{-HC}_{(2,1)}) \leq \mathrm{OPT}(k\text{-HC}_{(\infty,1)}) \leq \mathrm{OPT}(k\text{-HC}_{(2,1)}),$$

$$\frac{1}{n}\,\mathrm{OPT}(k\text{-HC}_{(2,1)}) \leq \mathrm{OPT}\Big(k\text{-HC}_{(1,\frac{1}{\sqrt{n}})}\Big) \leq \mathrm{OPT}(k\text{-HC}_{(2,1)}).$$

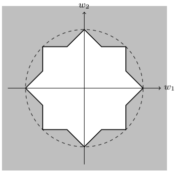
(A nontrivial consequence of norm equivalence)

Both polyhedral relaxations yield $\frac{1}{n}$-approximate solutions.

Multi-norm relaxation: a better $\frac{1}{\sqrt{n}}$-approximation

Impose *both* polyhedral constraints simultaneously:

$$\|w_j\|_1 \geq 1 \quad \text{and} \quad \|w_j\|_\infty \geq \frac{1}{\sqrt{n}}, \quad j \in [k].$$



k-HC$_{(multi,1)}$.

Theorem (congruence inequality)

Such a set induces a point-to-hyperplane distance stemming from the quantity $\max\left\{\|x\|_\infty, \frac{1}{\sqrt{n}}\|x\|_1\right\}$ *which 1) is a norm and 2) satisfies the congruence relationship*

$$n^{-\frac{1}{4}}\|x\|_2 \leq \max\left\{\|x\|_\infty, \frac{1}{\sqrt{n}}\|x\|_1\right\} \leq \|x\|_2.$$

Theorem

$$\frac{1}{\sqrt{n}}\,\mathrm{OPT}(k\text{-HC}_{(2,1)}) \leq \mathrm{OPT}(k\text{-HC}_{(\mathrm{multi},1)}) \leq \mathrm{OPT}(k\text{-HC}_{(2,1)}).$$

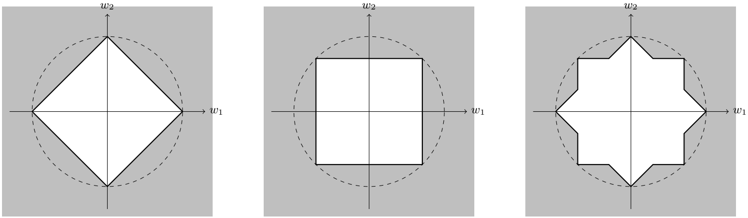
The multi-norm achieves a $\frac{1}{\sqrt{n}}$-approximation—strictly better than $\frac{1}{n}$.

Enhanced formulations

Rather than replacing the 2-norm, we can *strengthen* it by imposing the $1/\infty$/multi constraints alongside $\|w_j\|_2 \geq 1$.

Enhanced formulations

$$\text{original 2-norm formulation} \;+\; \|w_j\|_1 \geq 1 \;\text{ and/or }\; \|w_j\|_\infty \geq \tfrac{1}{\sqrt{n}}, \quad j \in [k].$$



k-HC$_{(\infty,1)})$ $\quad$ k-HC$_{(1,\frac{1}{\sqrt{n}})}$ $\quad$ k-HC$_{(multi,1)}$.

Proposition

With ∞-norm constraints, a nonzero global lower bound is obtained after at most $k(n-1)$ branching nodes—exponentially fewer than the baseline.

Relaxation	Approx. factor	Nodes for nonzero LB
1-norm only	$\frac{1}{n}$	$\geq 2^{k(n-1)}$
scaled ∞-norm only	$\frac{1}{n}$	$\leq k(n-1)$
multi-norm (both together)	$\frac{1}{\sqrt{n}}$	$\leq k(n-1)$

Experimental setup

Solver & hardware

Gurobi 10 with SBB; 12 threads; 2.6 GHz Intel Core i7-9750H; 32 GB RAM; time limit: 168,000 s (46 h).

`Low-dim` testbed (43 instances)

$m = 10, \ldots, 30; \quad n = 2, 3; \quad k = 2, 3$

(Superset of the 24 instances in Amaldi & Coniglio, EJOR 2013)

`High-dim` testbed (43 instances)

$m = 10, \ldots, 17; \quad n = 2, 3, 4, 5; \quad k = 2, 3, 4, 5$

Computational results: summary tables

`Low-dim`: 20 instances solved by all four.

Algorithm	**Med. (s)**	**Speed-up**	**p-val.**
(k-HC$_{(2,1)}$)	169.9	$1\times$	–
(k-HC$_{(2,1),(1,\frac{1}{\sqrt{n}})}$)	4.15	$40.9\times$	1.1×10^{-4}
(k-HC$_{(2,1),(\infty,1)}$)	6.10	$27.9\times$	1.1×10^{-4}
(k-HC$_{(\text{multi},1)}$)	5.00	$34.0\times$	1.1×10^{-4}

`High-dim`: 30 instances solved by all four.

Algorithm	**Med. (s)**	**Speed-up**	**p-val.**
(k-HC$_{(2,1)}$)	208.6	$1\times$	–
(k-HC$_{(2,1),(1,\frac{1}{\sqrt{n}})}$)	18.20	$11.5\times$	5.6×10^{-9}
(k-HC$_{(2,1),(\infty,1)}$)	20.65	$10.1\times$	7.5×10^{-9}
(k-HC$_{(\text{multi},1)}$)	37.35	$5.6\times$	8.7×10^{-4}

Wilcoxon signed-rank test, Holm-corrected ($\alpha = 0.05$).

Up to **41× speedup**; up to **63% more instances solved**.

Thanks for the attention!