Recall : Finite set $S \subseteq \mathbb{R}^n$

$p \in S$

$$V(p) = \{ v \in \mathbb{R}^n : \|p - v\| \leq \|p' - v\| \text{ for all } p' \in P \}.$$

Recall $u = (u_1, \ldots, u_n)$, $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$

$$u \cdot v = u v^t$$

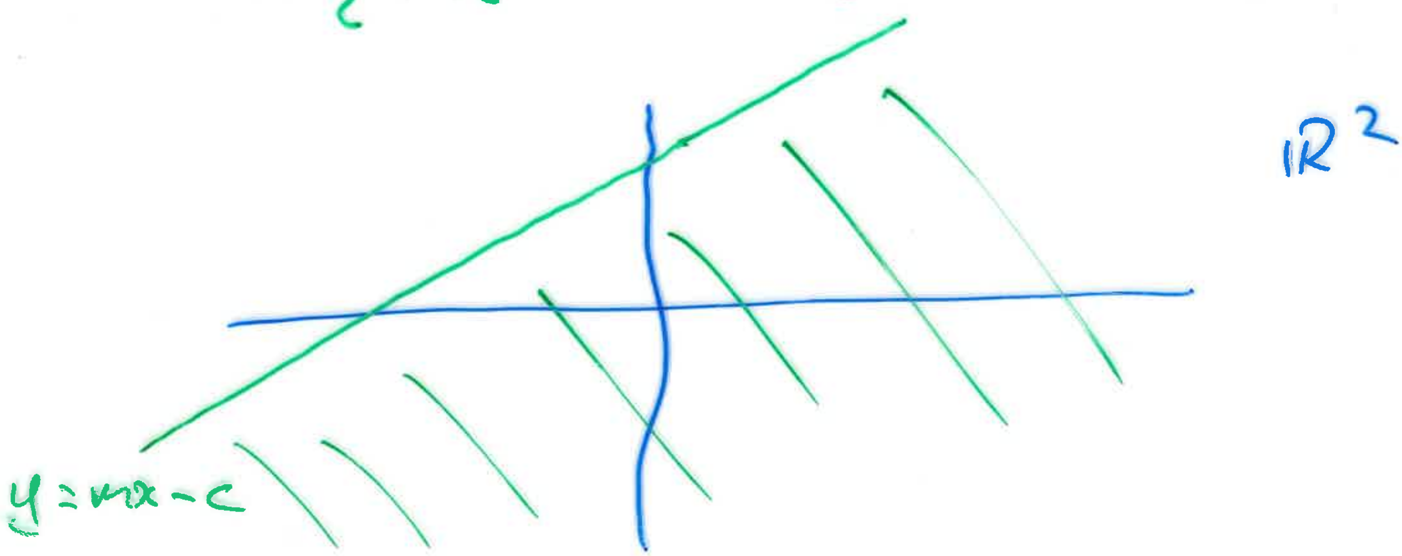
Defn For $w \in \mathbb{R}^n$ and any real number c, we define the half-space

$$H(w, c) = \{ v \in \mathbb{R}^n : w \cdot v \geq c \}$$

Example For $w = (m, -1) \in \mathbb{R}^2$, $c \in \mathbb{R}$

$$H(w, c) = \{ (x, y) \in \mathbb{R}^2 : (x, y) \cdot (m, -1) \geq c \}$$

$$= \{ (x, y) \in \mathbb{R}^2 : y \leq mx - c \}$$



Theorem The Voronoi region $V(p)$ is an intersection of finitely many half-spaces

$$V(p) = H(w_1, c_1) \cap H(w_2, c_2) \cap \dots \cap H(w_k, c_k)$$

In this theorem we can choose half spaces such that

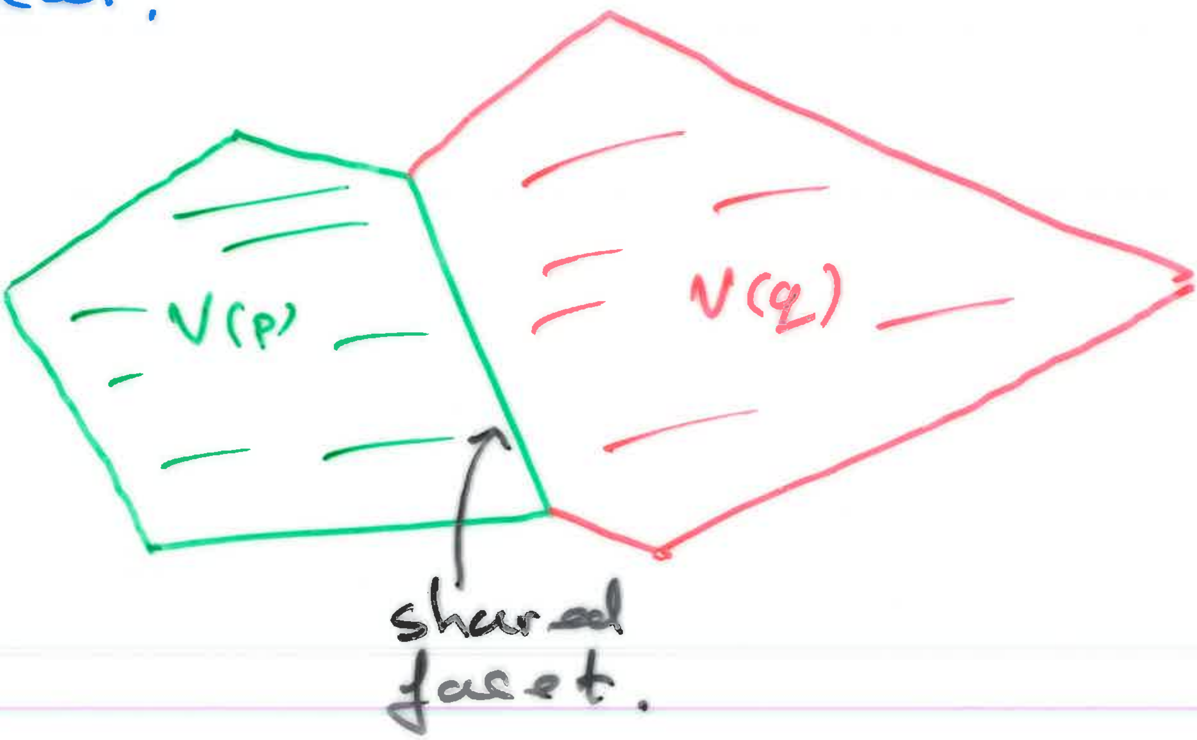
$$F_i = V(p) \cap H(-w_i, c_i)$$

is an $(n-1)$-dimensional subset of $\mathbb{R}^n$. (i.e. contains n vectors $v_0, v_1, \dots, v_{n-1} \in \mathbb{R}^n$ such

$$v_1 - v_0,\ v_2 - v_0,\ \dots,\ v_{n-1} - v_0$$

are linearly independent.) Those F_i are called facets.

Defn we say that points $p, q \in S$ are neighbours if their Voronoi regions share a facet.



For our database of points $S \subseteq \mathbb{R}^n$ we compute $V(p)$ for each $p \in S$, and then record for each p the neighbours

$$N(p) = \left\{ q \in S : \begin{array}{l} q \neq p \text{ and } V(p) \\ \text{shares a facet} \\ \text{with } V(q) \end{array} \right\}.$$

Algorithm

Input: $S \subset \mathbb{R}^n$ and $v \in \mathbb{R}^n$

Output: The point p in the database closest to v.

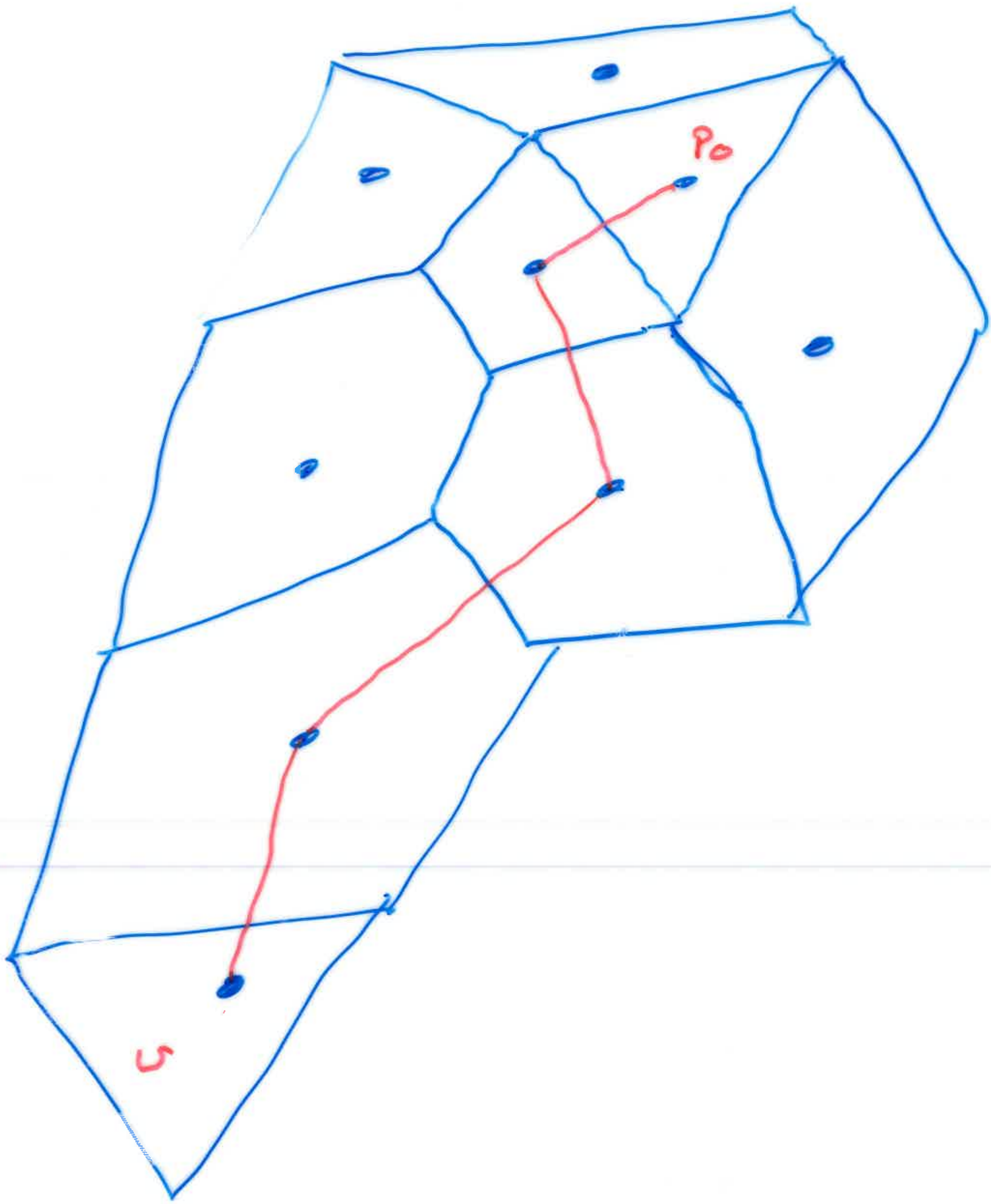
Procedure

1. Choose some $p_0 \in S$

2. while $\exists\ q \in N(p_0)$ with $\|q - v\| < \|p_0 - v\|$ do

   set $p_0 := q$

   end while;



p0
s

Theorem (Johnson–Lindenstrauss)

Let $\varepsilon \in (0, \frac{1}{2})$. Let $S \subseteq \mathbb{R}^d$ be a set of n points. Set

$$k = \frac{20 \log(n)}{\varepsilon^2}$$

There exists a linear transformation $f : \mathbb{R}^d \longrightarrow \mathbb{R}^k$ such that for all $u, v \in S$

$$(1-\varepsilon)\|u-v\|^2 \leq \|f(u)-f(v)\|^2 \leq (1+\varepsilon)\|u-v\|^2$$

Example If $S \subseteq \mathbb{R}^{1000000}$ has $n = |S| = 10000$ and if $\varepsilon = 0.499$

then $k = 20 \cdot 4 \cdot 4 = 320$

$$f : \mathbb{R}^{1000000} \longrightarrow \mathbb{R}^{320}$$

$$0.7\|u-v\| \leq \|f(u)-f(v)\| \leq 1.23\|u-v\|$$

Proof of J-L theorem depends on:

<u>Proposition</u> (Norm Preservation)

Let $x \in \mathbb{R}^d$. Assume that A is a $k \times d$ matrix whose kd entries are sampled independently from $N(0,1)$. Then

$$\Pr\left((1-\varepsilon)\|x\|^2 \leq \left\| \frac{1}{k} A x \right\| \leq (1+\varepsilon) \| x^2 \| \right)$$

$$\geq 1 - 2e^{-(\varepsilon^2 - \varepsilon^3)k/4}$$