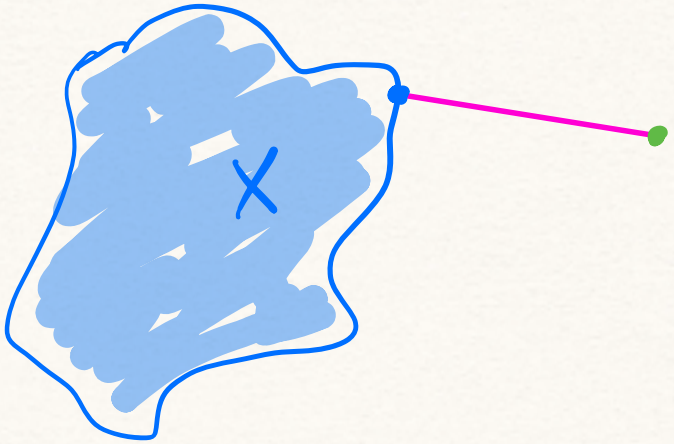
Algebraic Optimization Problems



Find closest point on an algebraic model from a given point.

Assumptions: (1) $X \subseteq \mathbb{R}^n$ is an irreducible algebraic variety

(2) Euclidean metric

~~(3) closest point is smooth on X~~

Today we challenge some of these assumptions!

Singular Stratification

$X \subseteq \mathbb{R}^n$ is an irreducible algebraic variety

Recall: The singular locus $\mathrm{Sing}(X)$ of X is a subvariety (i.e., Zariski closed).

Singular Stratification

$X \subseteq \mathbb{R}^n$ is an irreducible algebraic variety

Recall: The singular locus $\mathrm{Sing}(X)$ of X is a subvariety (i.e., Zariski closed).

It is in general not irreducible!

↳ Write $\mathrm{Sing}(X) = X_1 \cup X_2 \cup \dots \cup X_n$ decomposition into irreducible varieties such that $\forall i \neq j$: $X_i \not\subseteq X_j$

Singular Stratification

$X \subseteq \mathbb{R}^n$ is an irreducible algebraic variety

Recall: The singular locus $\mathrm{Sing}(X)$ of X is a subvariety (i.e., Zariski closed).

It is in general not irreducible!

↳ Write $\mathrm{Sing}(X) = X_1 \cup X_2 \cup \dots \cup X_n$ decomposition into irreducible varieties such that $\forall i \neq j$: $X_i \not\subseteq X_j$

$$\Rightarrow \mathrm{Sing}(\mathrm{Sing}(X)) = \mathrm{Sing}(X_1) \cup \dots \cup \mathrm{Sing}(X_n) \cup \bigcup_{i<j} (X_i \cap X_j)$$

Singular Stratification

$X \subseteq \mathbb{R}^n$ is an irreducible algebraic variety

Recall: The singular locus $\mathrm{Sing}(X)$ of X is a subvariety (i.e., Zariski closed).

It is in general not irreducible!

↳ Write $\mathrm{Sing}(X) = C_1 \cup C_2 \cup \dots \cup C_n$ decomposition into irreducible varieties such that $\forall i \neq j$: $C_i \not\subseteq C_j$

$$\Rightarrow \mathrm{Sing}(\mathrm{Sing}(X)) = \mathrm{Sing}(C_1) \cup \dots \cup \mathrm{Sing}(C_n) \cup \bigcup_{i<j} (C_i \cap C_j)$$

Iterate this process: Define $X_0 := X$, $X_i := \mathrm{Sing}(X_{i-1})$ for $i > 0$

Singular Stratification

$X \subseteq \mathbb{R}^n$ is an irreducible algebraic variety

Recall: The singular locus $\mathrm{Sing}(X)$ of X is a subvariety (i.e., Zariski closed).

It is in general not irreducible!

↳ Write $\mathrm{Sing}(X) = C_1 \cup C_2 \cup \dots \cup C_n$ decomposition into irreducible varieties such that $\forall i \neq j: C_i \not\subseteq C_j$

$$\Rightarrow \mathrm{Sing}(\mathrm{Sing}(X)) = \mathrm{Sing}(C_1) \cup \dots \cup \mathrm{Sing}(C_n) \cup \bigcup_{i<j} (C_i \cap C_j)$$

Iterate this process: Define $X_0 := X$, $X_i := \mathrm{Sing}(X_{i-1})$ for $i > 0$

$$\Rightarrow \exists k \leq n: X_k = \emptyset \quad \& \quad X = \bigcup_{i=0}^{k-1} \underbrace{X_i \setminus X_{i+1}}_{\text{smooth}}$$

Singular Stratification

$X \subseteq \mathbb{R}^n$ is an irreducible algebraic variety

Recall: The singular locus $\mathrm{Sing}(X)$ of X is a subvariety (i.e., Zariski closed).

It is in general not irreducible!

↳ Write $\mathrm{Sing}(X) = C_1 \cup C_2 \cup \dots \cup C_n$ decomposition into irreducible varieties such that $\forall i \neq j: C_i \not\subseteq C_j$

$$\Rightarrow \mathrm{Sing}(\mathrm{Sing}(X)) = \mathrm{Sing}(C_1) \cup \dots \cup \mathrm{Sing}(C_n) \cup \bigcup_{i<j} (C_i \cap C_j)$$

Iterate this process: Define $X_0 := X$, $X_i := \mathrm{Sing}(X_{i-1})$ for $i > 0$

$$\Rightarrow \exists k \leq n: X_k = \emptyset \quad \& \quad X = \bigcup_{i=0}^{k-1} \underbrace{X_i \setminus X_{i+1}}_{\text{smooth}}$$

We can perform the critical point computation on each irreducible component of each $X_i \setminus X_{i+1}$!

Example

```
i1 : QQ[x,y,z];

i2 : X0 = ideal(z^2-(1+x^2)*(y-1)^2*(y^2-x^3)^2);

o2 : Ideal of QQ[x..z]

i3 : X1 = radical ideal singularLocus X0;

o3 : Ideal of QQ[x..z]

i4 : decompose X1

o4 = {ideal(z, y-1), ideal (z, x^3-y^2)}

o4 : List

i5 : X2 = radical ideal singularLocus X1;

o5 : Ideal of QQ[x..z]

i6 : decompose X2

o6 = {ideal(z, y, x), ideal(z, y-1, x-1), ideal (z, y-1, x^2+x+1)}

o6 : List
```

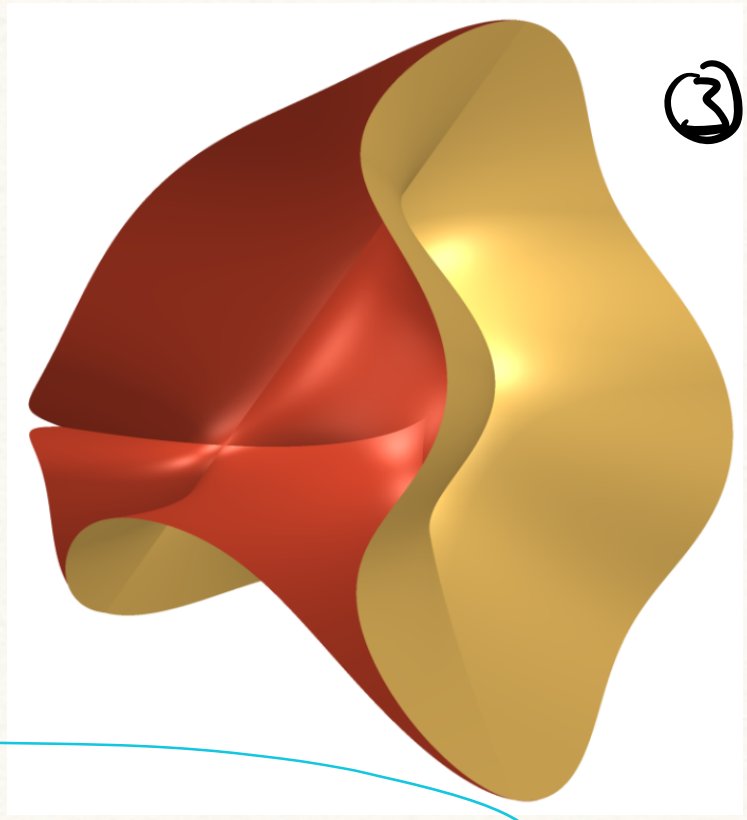
line L

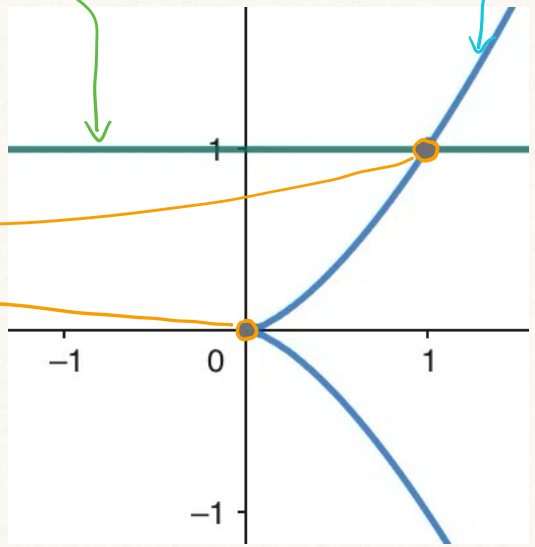
singular curve C

line & curve meet at 2 non-real points

P

Q





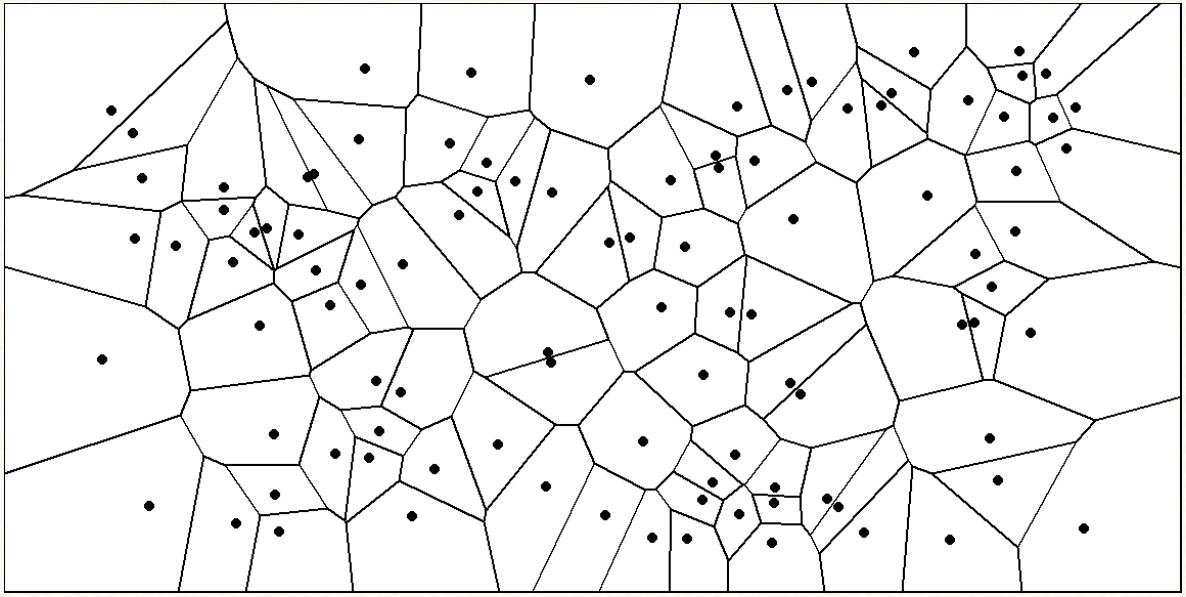
X_0 = disjoint union of smooth irreducible loci:

$X_0 \setminus (L \cup C)$	$L \setminus Q$	$C \setminus (P \cup Q)$	P	Q

To make sure that we find the global minimum, we can perform the critical points computation on each smooth irreducible locus along the singular stratification of X.

(When) is this necessary?

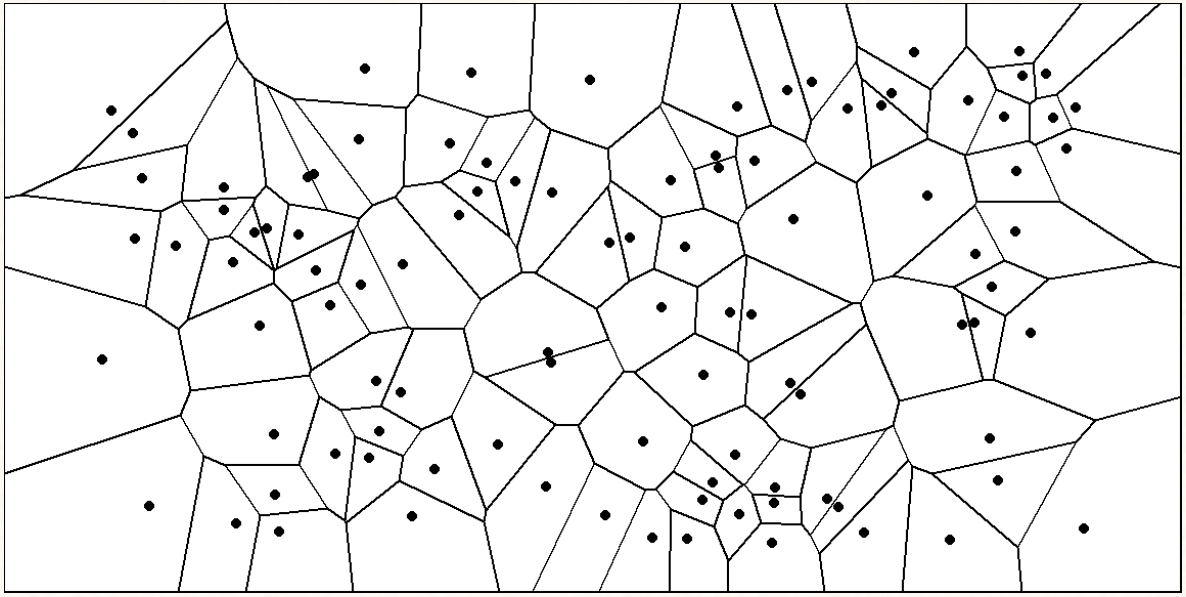
Voronoi cells



For $X \subseteq \mathbb{R}^n$,
the **Voronoi cell** at $p \in X$ is

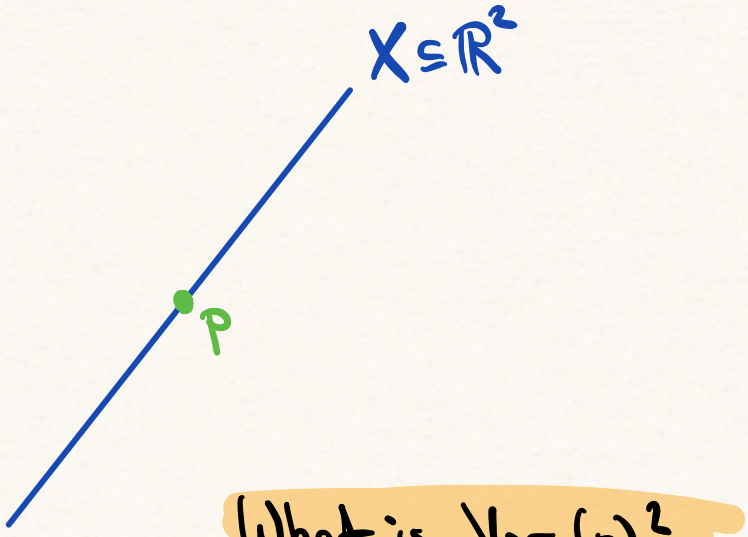
$$\mathrm{Vor}_X(p) := \{u \in \mathbb{R}^n \mid p \in \operatorname{arg\,min}_{x \in X} \|x-u\|^2\}.$$

Voronoi cells



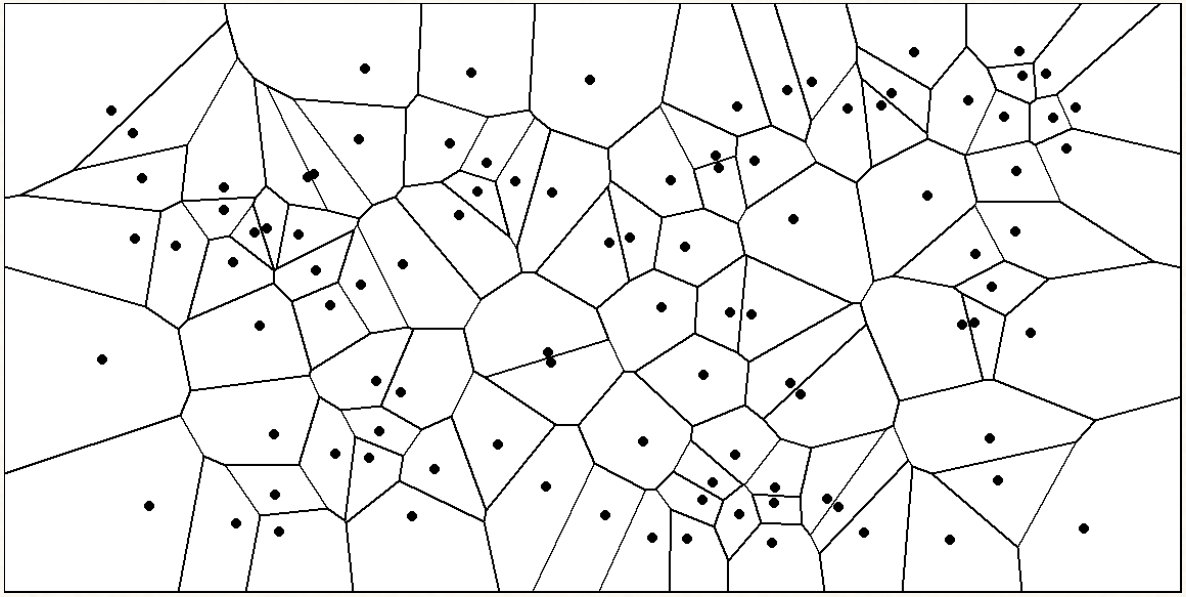
For $X \subseteq \mathbb{R}^n$,
the **Voronoi cell** at $p \in X$ is

$$\mathrm{Vor}_X(p) := \{u \in \mathbb{R}^n \mid p \in \arg\min_{x \in X} \|x-u\|^2\}.$$



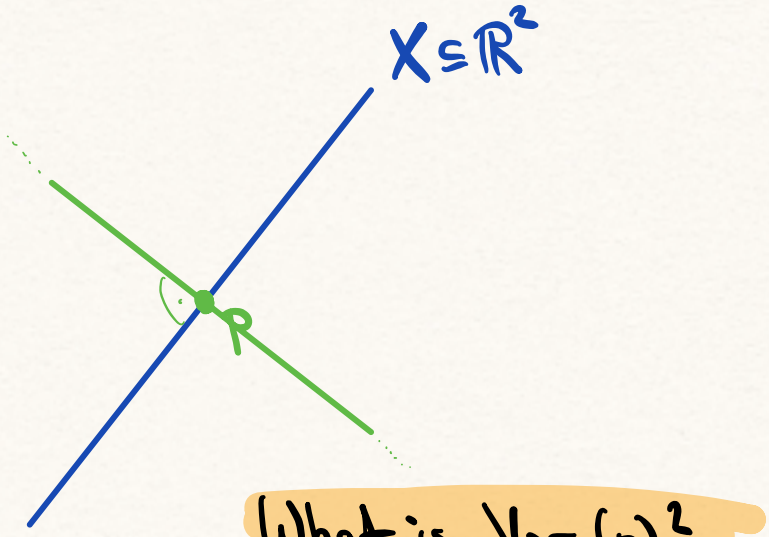
What is $\mathrm{Vor}_X(p)$?

Voronoi cells



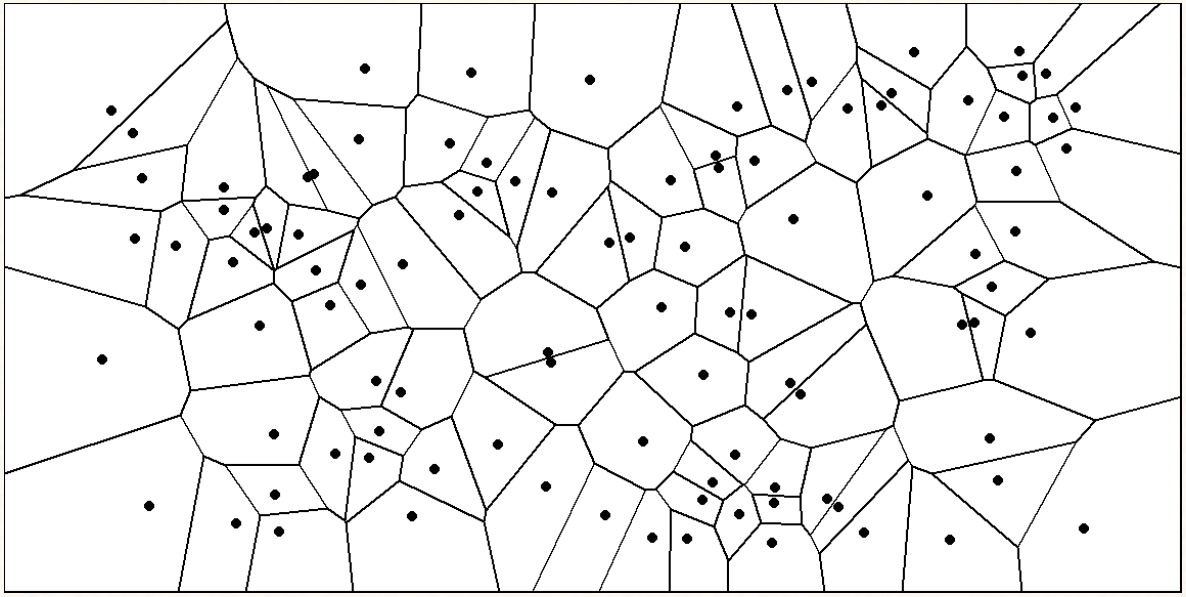
For $X \subseteq \mathbb{R}^n$,
the **Voronoi cell** at $p \in X$ is

$$\mathrm{Vor}_X(p) := \{u \in \mathbb{R}^n \mid p \in \arg\min_{x \in X} \|x-u\|^2\}.$$



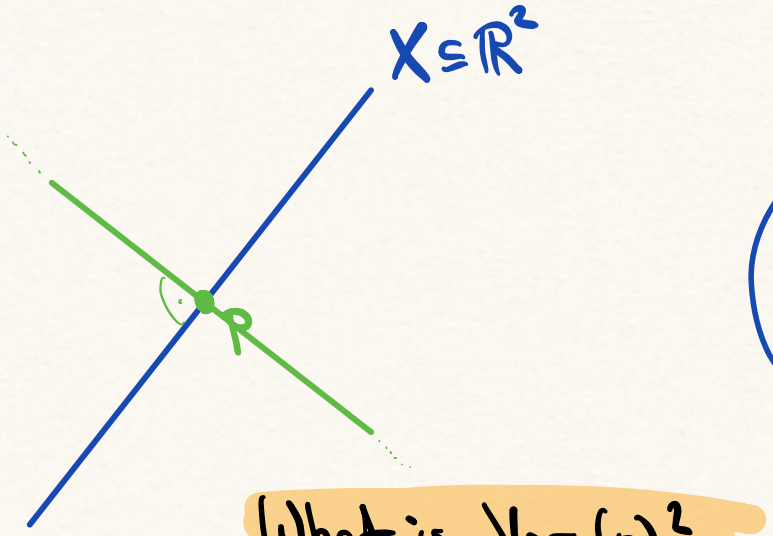
What is $\mathrm{Vor}_X(p)$?

Voronoi cells

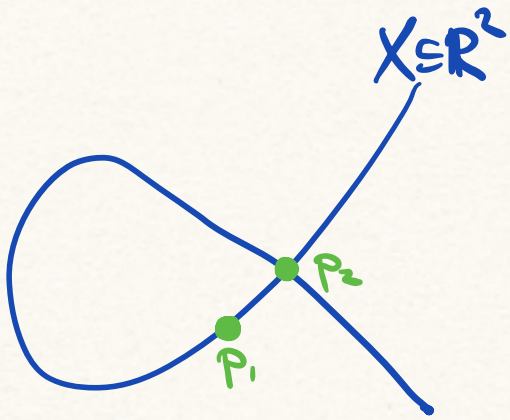


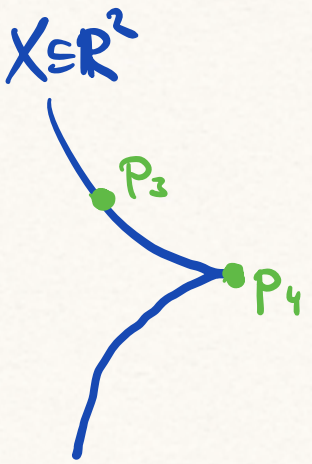
For $X \subseteq \mathbb{R}^n$,
the **Voronoi cell** at $p \in X$ is

$$\mathrm{Vor}_X(p) := \{u \in \mathbb{R}^n \mid p \in \arg\min_{x \in X} \|x-u\|^2\}.$$

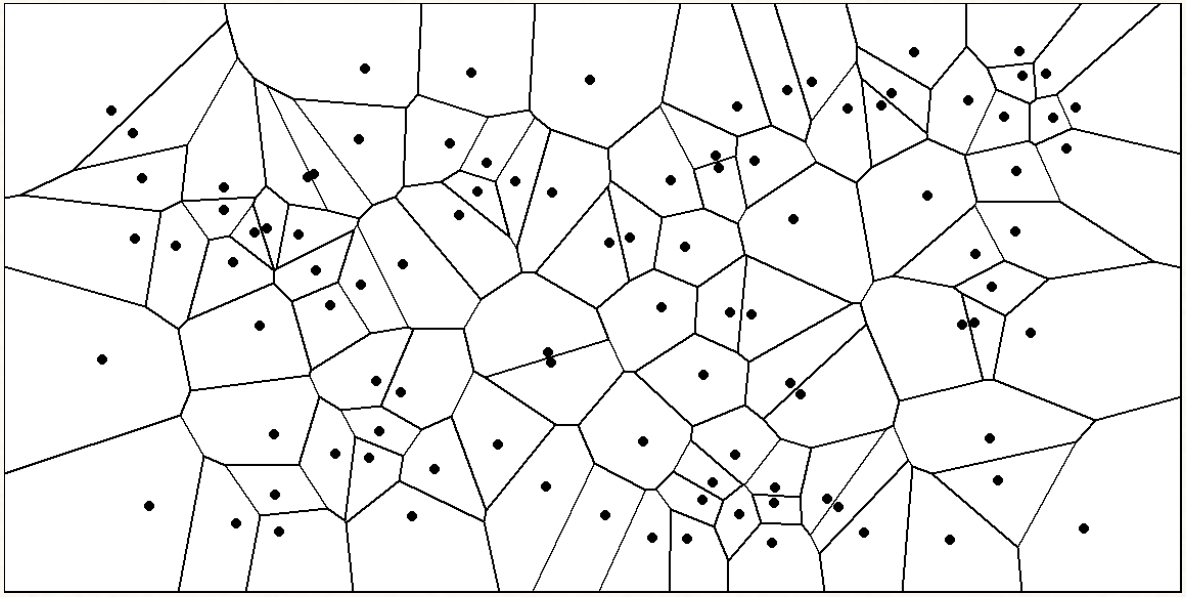


What is $\mathrm{Vor}_X(p)$?



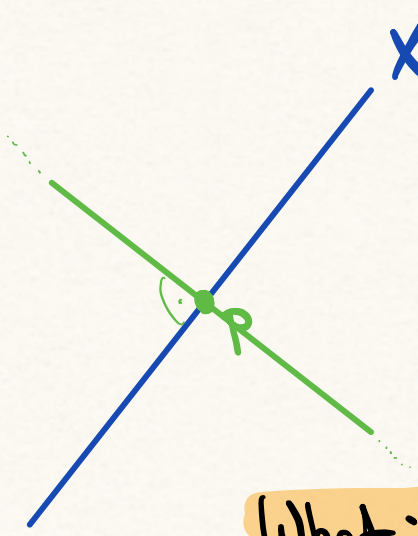


Voronoi cells

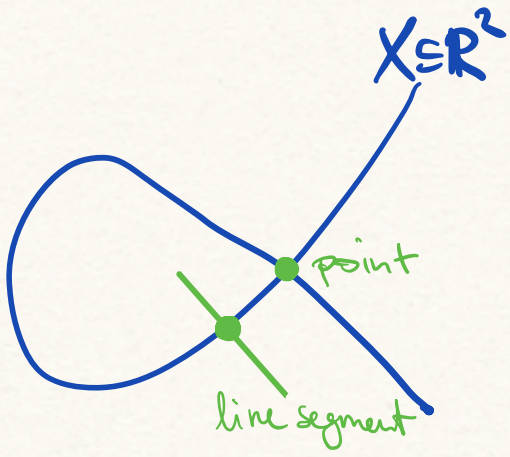


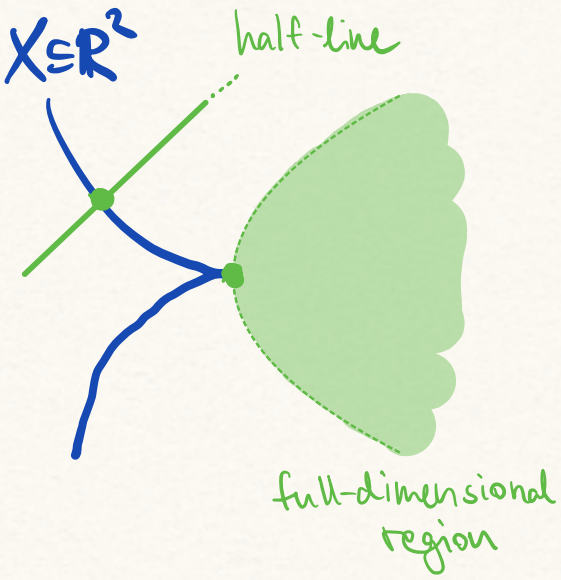
For $X \subseteq \mathbb{R}^n$,
the **Voronoi cell** at $p \in X$ is

$$\mathrm{Vor}_X(p) := \{u \in \mathbb{R}^n \mid p \in \arg\min_{x \in X} \|x-u\|^2\}.$$



What is $\mathrm{Vor}_X(p)$?





Voronoi cells at smooth points

Let $p \in X \subseteq \mathbb{R}^n$ be a smooth point.

Recall:
- p is a critical point of $\min_{x \in X} \|x-u\|^2$ iff $u \in \underbrace{p + (T_p X)^\perp}_{\text{affine normal space of } X \text{ at } p} = \{w \in \mathbb{R}^n \mid \forall v \in T_p X : w^T v = 0\}$
- $\dim(p + (T_p X)^\perp) = n - \dim X$

Voronoi cells at smooth points

Let $p \in X \subseteq \mathbb{R}^n$ be a smooth point.

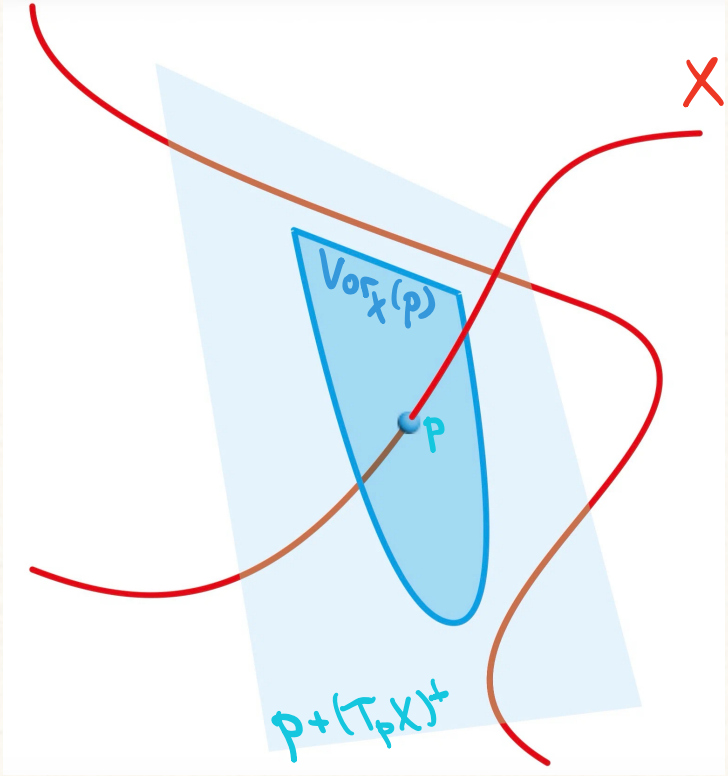
Recall: • p is a critical point of $\min_{x \in X} \|x-u\|^2$ iff $u \in p + \underbrace{(T_pX)^\perp}_{\text{affine normal space of } X \text{ at } p}$

$= \{w \in \mathbb{R}^n \mid \forall v \in T_pX : w^T v = 0\}$

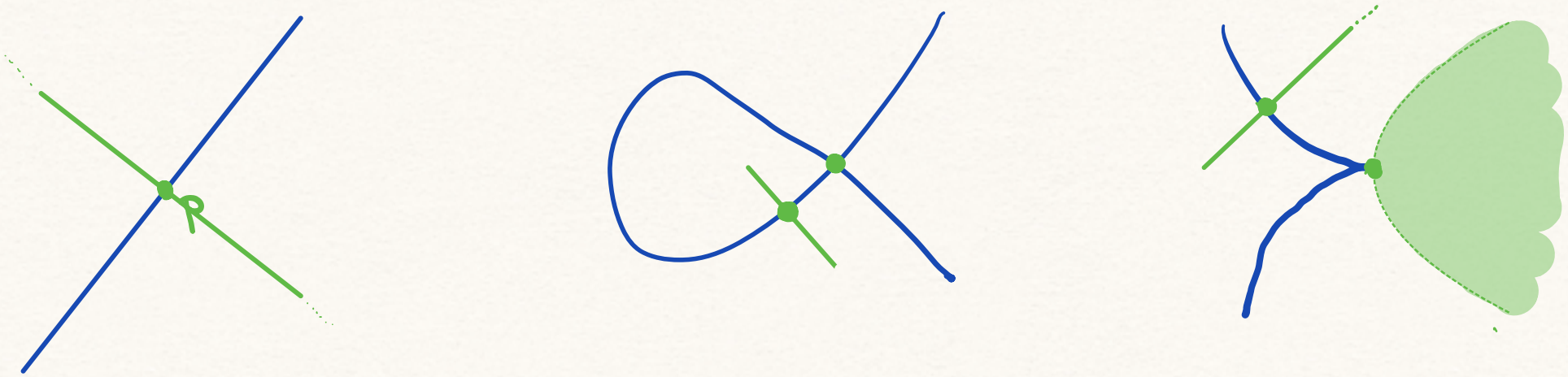
• $\dim(p+(T_pX)^\perp) = n - \dim X$

Proposition:
$\mathrm{Vor}_X(p)$ is a convex semialgebraic full-dimensional subset of $p+(T_pX)^\perp$.

In particular, at all smooth points, the Voronoi cell has dimension $n-\dim X$.



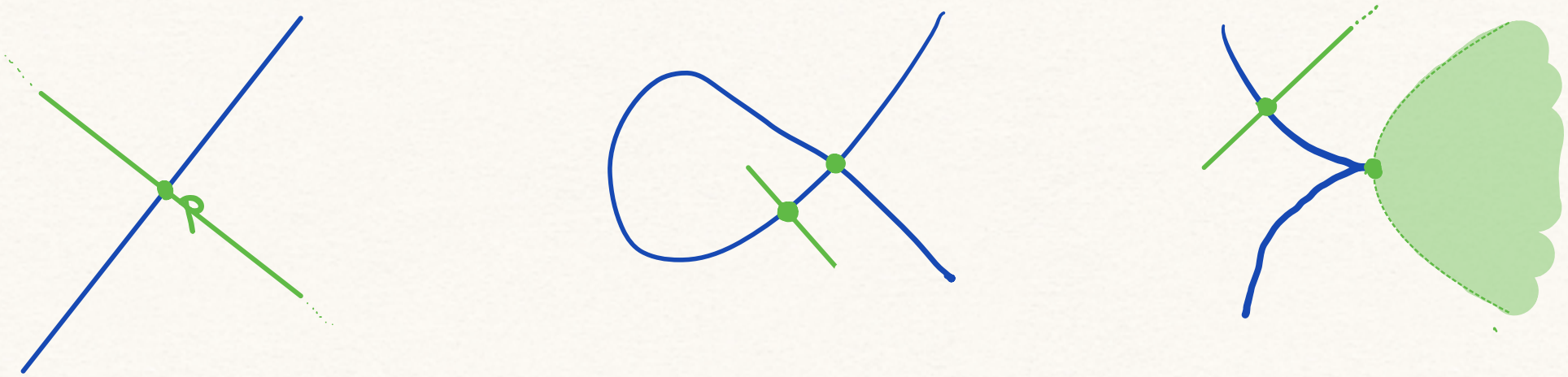
Voronoi cells at singular points



At singular points, the Voronoi cell dimension can be different from n-dim X:

It can be larger or smaller!

Voronoi cells at singular points



At singular points, the Voronoi cell dimension can be different from n-dim X:

It can be larger or smaller!

Definition: Let C be an irreducible component of $\mathrm{Sing}(X)$.

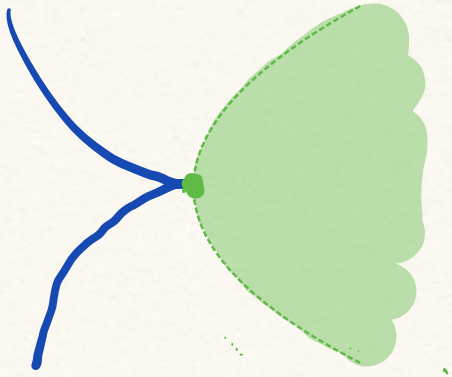
If $\dim\left(\bigcup_{p \in C} \mathrm{Vor}_X(p)\right) = n$, then C is **exposed** w.r.t.

Euclidean distance minimization to X.

Voronoi cells at singular points

Definition: Let C be an irreducible component of $\mathrm{Sing}(X)$.

If $\dim\left(\bigcup_{p \in C} \mathrm{Vor}_X(p)\right) = n$, then C is **exposed** w.r.t. Euclidean distance minimization to X.



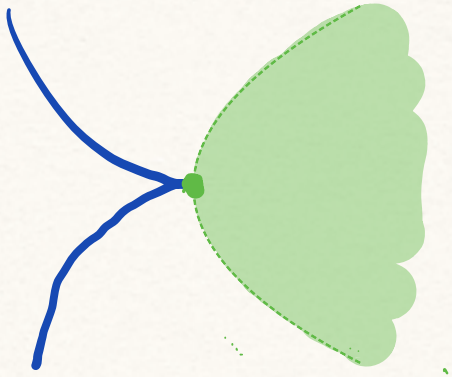
If C is exposed:

For all data distributions on $\mathbb{R}^n$ with full support, $\underset{x \in X}{\arg\min} \|x-u\|^2$ contains a point on C with positive probability on the data $u \in \mathbb{R}^n$.

Voronoi cells at singular points

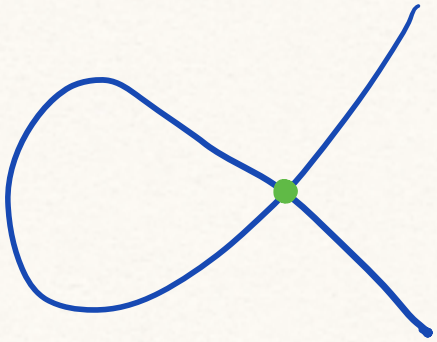
Definition: Let C be an irreducible component of $\mathrm{Sing}(X)$.

If $\dim\left(\bigcup_{p \in C} \mathrm{Vor}_X(p)\right) = n$, then C is **exposed** w.r.t. Euclidean distance minimization to X.



If C is exposed:

For all data distributions on $\mathbb{R}^n$ with full support, $\underset{x \in X}{\operatorname{argmin}} \|x-u\|^2$ contains a point on C with positive probability on the data $u \in \mathbb{R}^n$.



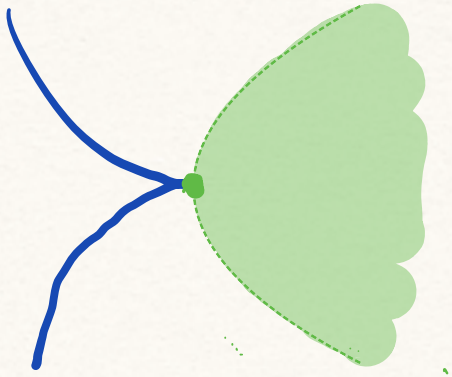
If C is not exposed:

For almost all $u \in \mathbb{R}^n$, $\underset{x \in X}{\operatorname{argmin}} \|x-u\|^2 \subseteq X \setminus C$.

Voronoi cells at singular points

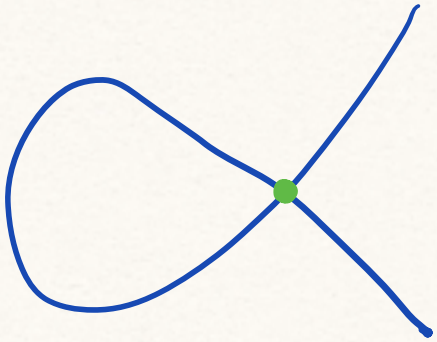
Definition: Let C be an irreducible component of $\mathrm{Sing}(X)$.

If $\dim\left(\bigcup_{p \in C} \mathrm{Vor}_X(p)\right) = n$, then C is **exposed** w.r.t. Euclidean distance minimization to X.



If C is exposed:

For all data distributions on $\mathbb{R}^n$ with full support, $\operatorname{argmin}_{x \in X} \|x-u\|^2$ contains a point on C with positive probability on the data $u \in \mathbb{R}^n$.



If C is not exposed:

For almost all $u \in \mathbb{R}^n$, $\operatorname{argmin}_{x \in X} \|x-u\|^2 \subseteq X \setminus C$.

Hence: For noisy data $u \in \mathbb{R}^n$, we can neglect C when solving $\min_{x \in X} \|x-u\|^2$ & don't have to compute critical points on C!

Recall: Low-Rank Matrix Approximation

$$X = \{ A \in \mathbb{R}^{m \times n} \mid \operatorname{rank}(A) \leq r \}$$

Given $U \in \mathbb{R}^{m \times n}$, solve: $\underset{A \in X}{\operatorname{argmin}} \|A - U\|_F^2$ ← Frobenius norm

$$= \sum_{i,j} (a_{ij} - u_{ij})^2$$

Assume: $m \leq n$

Recall: A singular value decomposition of U is

$$U = T_1 \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_m \end{bmatrix} T_2, \quad \text{where } T_1 \in O(m),\ T_2 \in O(n),$$

$\sigma_1 \geq \sigma_2 \geq \dots \sigma_m \geq 0$.

singular values of U

Eckart-Young Theorem:

A closest matrix on X to U is $T_1 \begin{bmatrix} \sigma_1 & & 0 \\ & \sigma_r & \\ 0 & & 0 \end{bmatrix} T_2$.

For almost all U, this is the unique local minimum.

All complex critical points are real: they are found by substituting 0 for $m-r$ entries of $\begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_m \end{bmatrix}$.

$EDD(X) = \binom{m}{r}$

Low-Rank Matrix Approximation - Voronoi cells

$$X = \{ A \in \mathbb{R}^{m \times n} \mid \operatorname{rank}(A) \leq r \}, \quad 0 < r < m \leq n$$

Fact: $\operatorname{Sing}(X) = \{ A \in \mathbb{R}^{m \times n} \mid \operatorname{rank}(A) \leq r-1 \}$

Low-Rank Matrix Approximation - Voronoi cells

$$X = \{ A \in \mathbb{R}^{m \times n} \mid \operatorname{rank}(A) \leq r \}, \quad 0 < r < m \leq n$$

Fact: $\operatorname{Sing}(X) = \{ A \in \mathbb{R}^{m \times n} \mid \operatorname{rank}(A) \leq r-1 \}$

Eckart-Young Theorem: A closest matrix on X to $T_1 \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_m \end{bmatrix} T_2$ ($\sigma_1 \geq \sigma_2 \geq \dots \sigma_m \geq 0$) is $T_1 \begin{bmatrix} \sigma_1 & & 0 \\ & \sigma_r & \\ 0 & & 0 \end{bmatrix} T_2$.

Given $A \in \operatorname{Sing}(X)$, what is $\operatorname{Vor}_X(A)$?

Low-Rank Matrix Approximation - Voronoi cells

$$X = \{ A \in \mathbb{R}^{m \times n} \mid \operatorname{rank}(A) \leq r \}, \quad 0 < r < m \leq n$$

Fact: $\operatorname{Sing}(X) = \{ A \in \mathbb{R}^{m \times n} \mid \operatorname{rank}(A) \leq r-1 \}$

Eckart-Young Theorem: A closest matrix on X to $T_1 \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_m \end{bmatrix} T_2$ ($\sigma_1 \geq \sigma_2 \geq \cdots \sigma_m \geq 0$) is $T_1 \begin{bmatrix} \sigma_1 & & 0 \\ & \sigma_r & \\ 0 & & 0 \end{bmatrix} T_2$.

Corollary: 1) For $A \in \operatorname{Sing}(X)$, $\operatorname{Vor}_X(A) = \{A\}$.

Low-Rank Matrix Approximation - Voronoi cells

$$X = \{ A \in \mathbb{R}^{m \times n} \mid \operatorname{rank}(A) \leq r \}, \quad 0 < r < m \leq n$$

Fact: $\operatorname{Sing}(X) = \{ A \in \mathbb{R}^{m \times n} \mid \operatorname{rank}(A) \leq r-1 \}$

Eckart-Young Theorem: A closest matrix on X to $T_1 \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_m \end{bmatrix} T_2$ $(\sigma_1 \geq \sigma_2 \geq \dots \sigma_m \geq 0)$ is $T_1 \begin{bmatrix} \sigma_1 & & 0 \\ & \sigma_r & \\ 0 & & 0 \end{bmatrix} T_2$.

Corollary: 1) For $A \in \operatorname{Sing}(X)$, $\operatorname{Vor}_X(A) = \{A\}$. 2) $\operatorname{Sing}(X)$ is unexposed.

Low-Rank Matrix Approximation - Voronoi cells

$$X = \{ A \in \mathbb{R}^{m \times n} \mid \operatorname{rank}(A) \leq r \}, \quad 0 < r < m \leq n$$

Fact: $\operatorname{Sing}(X) = \{ A \in \mathbb{R}^{m \times n} \mid \operatorname{rank}(A) \leq r-1 \}$

Eckart-Young Theorem: A closest matrix on X to $T_1 \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_m \end{bmatrix} T_2$ ($\sigma_1 \geq \sigma_2 \geq \dots \sigma_m \geq 0$) is $T_1 \begin{bmatrix} \sigma_1 & & 0 \\ & \sigma_r & \\ 0 & & 0 \end{bmatrix} T_2$.

Corollary: 1) For $A \in \operatorname{Sing}(X)$, $\operatorname{Vor}_X(A) = \{A\}$. 2) $\operatorname{Sing}(X)$ is unexposed.

Proposition: For $A = T_1 \begin{bmatrix} \sigma_1 & & 0 \\ & \sigma_r & \\ 0 & & 0 \end{bmatrix} T_2$ of rank r,

$$\operatorname{Vor}_X(A) = \left\{ T_1 \left[\begin{array}{ccc|c} \sigma_1 & & & 0 \\ & \ddots & & \\ & & \sigma_r & \\ \hline & 0 & & E \end{array}\right] T_2 \;\middle|\; E \in \mathbb{R}^{(m-r)\times(n-r)}, \; \|E\|_{op} \leq \sigma_r \right\}.$$

ball of radius σ_r in spectral norm

Proof of Proposition

$$\mathrm{Vor}_X(A) \subseteq A + (T_A X)^{\perp}$$

$\perp$ ← w.r.t. Frobenius inner product $\langle M, N \rangle_F := \mathrm{tr}(M^T N)$

Proof of Proposition

$$\mathrm{Vor}_X(A) \subseteq A + (T_A X)^{\perp}$$

$\perp$ ← w.r.t. Frobenius inner product $\langle M, N \rangle_F := \mathrm{tr}(M^T N)$

Exercise: For $A = T_1 \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \\ & 0 & \end{bmatrix} T_2$:

$$T_A X = \left\{ T_1 \begin{bmatrix} B & C \\ D & 0 \end{bmatrix} T_2 \;\middle|\; \begin{array}{l} B \in \mathbb{R}^{r \times r},\ C \in \mathbb{R}^{r \times (n-r)} \\ D \in \mathbb{R}^{(m-r) \times r} \end{array} \right\}$$

Proof of Proposition

$$\mathrm{Vor}_X(A) \subseteq A + (T_A X)^{\perp}$$

$\perp$ ← w.r.t. Frobenius inner product $\langle M, N \rangle_F := \mathrm{tr}(M^T N)$

Exercise: For $A = T_1 \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \\ & 0 & \end{bmatrix} T_2$:

$$T_A X = \left\{ T_1 \begin{bmatrix} B & C \\ D & 0 \end{bmatrix} T_2 \;\middle|\; \begin{matrix} B \in \mathbb{R}^{r \times r}, C \in \mathbb{R}^{r \times (n-r)}, \\ D \in \mathbb{R}^{(m-r) \times r} \end{matrix} \right\} \quad \& \quad (T_A X)^{\perp} = \left\{ T_1 \begin{bmatrix} 0 & 0 \\ 0 & E \end{bmatrix} T_2 \mid E \in \mathbb{R}^{(m-r) \times (n-r)} \right\}.$$

Proof of Proposition

$$\mathrm{Vor}_X(A) \subseteq A + (T_A X)^{\perp}$$ ← w.r.t. Frobenius inner product $\langle M, N \rangle_F := \mathrm{tr}(M^T N)$

Exercise: For $A = T_1 \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \\ & 0 & \end{bmatrix} T_2$:

$$T_A X = \left\{ T_1 \begin{bmatrix} B & C \\ D & 0 \end{bmatrix} T_2 \;\middle|\; \begin{matrix} B \in \mathbb{R}^{r \times r}, C \in \mathbb{R}^{r \times (n-r)}, \\ D \in \mathbb{R}^{(m-r) \times r} \end{matrix} \right\} \quad \& \quad (T_A X)^{\perp} = \left\{ T_1 \begin{bmatrix} 0 & 0 \\ 0 & E \end{bmatrix} T_2 \;\middle|\; E \in \mathbb{R}^{(m-r) \times (n-r)} \right\}.$$

Hence: $$\mathrm{Vor}_X(A) \subseteq \left\{ T_1 \left[\begin{array}{ccc|c} \sigma_1 & & & 0 \\ & \ddots & & \\ & & \sigma_r & \\ \hline 0 & & & E \end{array}\right] T_2 \;\middle|\; E \in \mathbb{R}^{(m-r) \times (n-r)} \right\}.$$

Proof of Proposition

$$\mathrm{Vor}_X(A) \subseteq A + (T_A X)^{\perp}$$ ← w.r.t. Frobenius inner product $\langle M, N \rangle_F := \mathrm{tr}(M^T N)$

Exercise: For $A = T_1 \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \\ & 0 & \end{bmatrix} T_2$:

$$T_A X = \left\{ T_1 \begin{bmatrix} B & C \\ D & 0 \end{bmatrix} T_2 \;\middle|\; \begin{matrix} B \in \mathbb{R}^{r \times r}, C \in \mathbb{R}^{r \times (n-r)} \\ D \in \mathbb{R}^{(m-r) \times r} \end{matrix} \right\} \quad \& \quad (T_A X)^{\perp} = \left\{ T_1 \begin{bmatrix} 0 & 0 \\ 0 & E \end{bmatrix} T_2 \;\middle|\; E \in \mathbb{R}^{(m-r) \times (n-r)} \right\}.$$

Hence: $$\mathrm{Vor}_X(A) \subseteq \left\{ T_1 \left[\begin{array}{ccc|c} \sigma_1 & & & 0 \\ & \ddots & & \\ & & \sigma_r & \\ \hline 0 & & & E \end{array}\right] T_2 \;\middle|\; E \in \mathbb{R}^{(m-r) \times (n-r)} \right\}.$$

By Eckart–Young Theorem: A is a closest point to this on X iff every singular value of E is at most σ_r, i.e., $\|E\|_{op} \leq \sigma_r$.

□

Example: $m=n=3$, $r=1$

Proposition: For $A = T_1 \begin{bmatrix} \sigma & 0 & 0 \\ 0 & & \\ 0 & & \end{bmatrix} T_2$ with $\sigma > 0$,

$$\mathrm{Vor}_X(A) = \left\{ T_1 \begin{bmatrix} \sigma & 0 & 0 \\ 0 & & \\ 0 & & E \end{bmatrix} T_2 \;\middle|\; E \in \mathbb{R}^{2\times 2},\ \|E\|_{op} \le \sigma \right\}.$$

ball of radius σ in spectral norm on $\mathbb{R}^{2\times 2}$

Example: $m=n=3$, $r=1$

Proposition: For $A = T_1 \begin{bmatrix} \sigma & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} T_2$ with $\sigma \geq 0$,

$$Vor_X(A) = \left\{ T_1 \begin{bmatrix} \sigma & 0 & 0 \\ 0 & & \\ 0 & & E \end{bmatrix} T_2 \;\middle|\; E \in \mathbb{R}^{2\times 2},\ \|E\|_{op} \leq \sigma \right\}.$$

ball of radius σ in spectral norm on $\mathbb{R}^{2\times 2}$

Restricting to symmetric matrices $X = \{ A \in \mathbb{R}^{m\times m} \mid A^T = A,\ \operatorname{rank}(A) \leq r \}$,

for $A = Q \begin{bmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \lambda_r & \\ & & & 0 \end{bmatrix} Q^T$ with $|\lambda_1| \geq \ldots \geq |\lambda_r| > 0$ (and $Q \in O(m)$),

we get the analog **Proposition:** $Vor_X(A) = \left\{ Q \begin{bmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \lambda_r & \\ & & & E \end{bmatrix} Q^T \;\middle|\; E^T = E,\ \|E\|_{op} \leq |\lambda_r| \right\}.$

Example: $m=n=3$, $r=1$

Proposition: For $A = T_1 \begin{bmatrix} \sigma & 0 \\ 0 & 0 \end{bmatrix} T_2$ with $\sigma \geq 0$,

$$\mathrm{Vor}_X(A) = \left\{ T_1 \begin{bmatrix} \sigma & 0 & 0 \\ 0 & & \\ 0 & & E \end{bmatrix} T_2 \;\middle|\; E \in \mathbb{R}^{2\times 2},\ \|E\|_{op} \leq \sigma \right\}.$$

ball of radius σ in spectral norm on $\mathbb{R}^{2\times 2}$

Restricting to symmetric matrices $X = \{ A \in \mathbb{R}^{m\times m} \mid A^T = A,\ \mathrm{rank}(A) \leq r \}$,

for $A = Q \begin{bmatrix} \lambda_1 & & & & \\ & \ddots & & & \\ & & \lambda_r & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{bmatrix} Q^T$ with $|\lambda_1| \geq \ldots \geq |\lambda_r| > 0$ (and $Q \in O(m)$),

we get the analog **Proposition:** $\mathrm{Vor}_X(A) = \left\{ Q \begin{bmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \lambda_r & \\ & & & E \end{bmatrix} Q^T \;\middle|\; E^T = E,\ \|E\|_{op} \leq |\lambda_r| \right\}$.

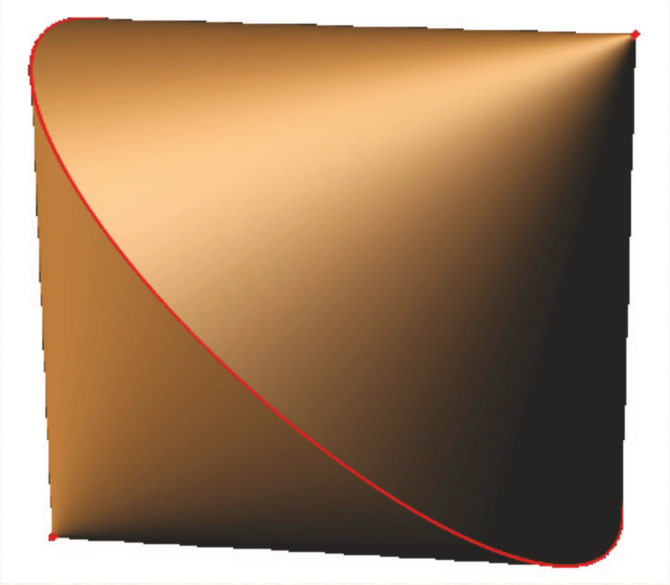
For $m=3$, $r=1$: $\mathrm{Vor}_X(A) = \left\{ Q \begin{bmatrix} \lambda & 0 & 0 \\ 0 & & \\ 0 & & E \end{bmatrix} Q^T \;\middle|\; E^T = E,\ \|E\|_{op} \leq |\lambda| \right\}$

ball of radius $|\lambda|$ in spectral norm on $\{E \in \mathbb{R}^{2\times 2} \mid E^T = E\} \cong \mathbb{R}^3$

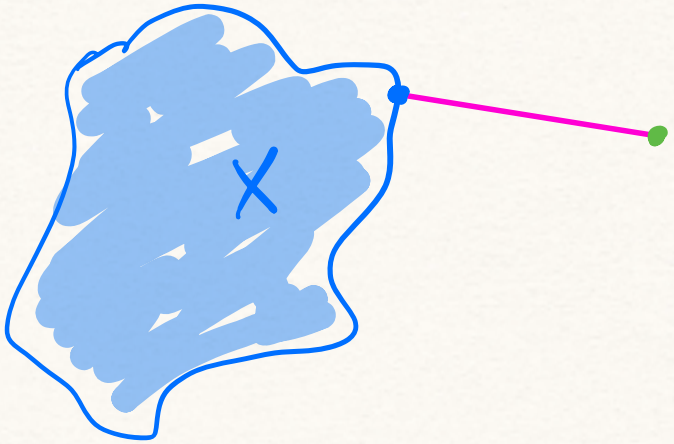
we can draw this :)

Example: $m=n=3$, $r=1$
symmetric matrices w.r.t. Frobenius norm

$$\mathrm{Var}_X\left(\begin{bmatrix}1&0\\0&0\end{bmatrix}\right) = \left\{\begin{bmatrix}1&0&0\\0&&\\0&&E\end{bmatrix} \;\middle|\; E^T=E,\ \|E\|_{op}\le 1\right\}$$



Algebraic Optimization Problems



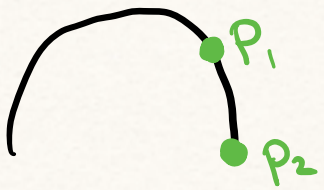
Find closest point on an algebraic model from a given point.

Assumptions:

1. $X \subseteq \mathbb{R}^n$ is ~~an irreducible algebraic variety~~ a semialgebraic set (whose Zariski closure is irreducible)
2. Euclidean metric
3. ~~closest point is smooth on X~~

Today we challenge some of these assumptions!

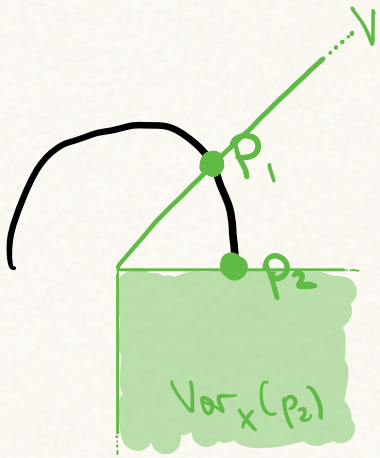
Voronoi cells of semialgebraic sets



$$X = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1,\ y \geq 0\}$$

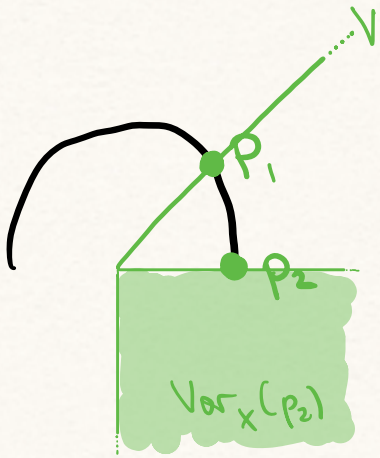
What are $\mathrm{Vor}_X(p_1)$ & $\mathrm{Vor}_X(p_2)$?

Voronoi cells of semialgebraic sets



$$X = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1,\ y \geq 0\}$$

Voronoi cells of semialgebraic sets

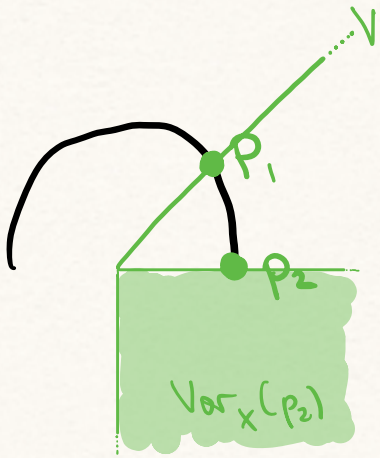


$$X = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1,\ y \geq 0\}$$



Are these 2 points the Euclidean boundary of X?

Voronoi cells of semialgebraic sets



$$X = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1,\ y \geq 0\}$$

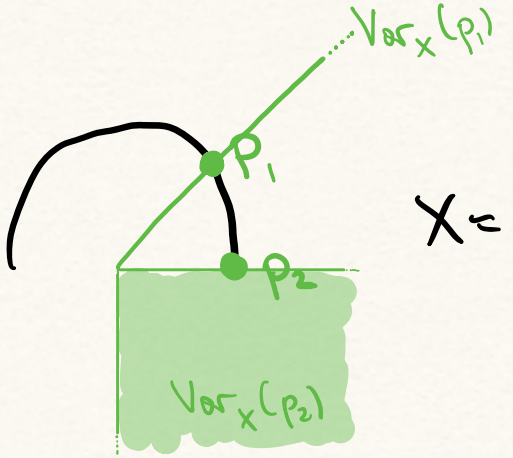


Are these 2 points the Euclidean boundary of X?

Not in the Euclidean topology of the ambient space $\mathbb{R}^2$!

↳ in that topology, every point of X lies on its boundary and so X has empty interior

Voronoi cells of semialgebraic sets



$$X = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1,\ y \geq 0\}$$



Are these 2 points the Euclidean boundary of X?

Not in the Euclidean topology of the ambient space $\mathbb{R}^2$!

↳ in that topology, every point of X lies on its boundary and so X has empty interior

YES in the Euclidean topology that the Zariski closure of X inherits from $\mathbb{R}^2$.

☺

Relative boundary

Let $X \subseteq \mathbb{R}^n$ be a semialgebraic set and let $\overline{X}^{Zar}$ be its Zariski closure inside $\mathbb{R}^n$.

Equip $\overline{X}^{Zar}$ with the subspace topology induced by the Euclidean topology on $\mathbb{R}^n$.

Relative boundary

Let $X \subseteq \mathbb{R}^n$ be a semialgebraic set and let $\overline{X}^{Zar}$ be its Zariski closure inside $\mathbb{R}^n$.

Equip $\overline{X}^{Zar}$ with the subspace topology induced by the Euclidean topology on $\mathbb{R}^n$.

Definition: The **relative boundary** of X inside $\overline{X}^{Zar}$, denoted by ∂X, is the boundary of X in the Euclidean subspace topology on $\overline{X}^{Zar}$.

Relative boundary

Let $X \subseteq \mathbb{R}^n$ be a semialgebraic set and let $\overline{X}^{Zar}$ be its Zariski closure inside $\mathbb{R}^n$.

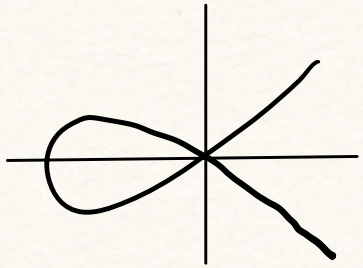
Equip $\overline{X}^{Zar}$ with the subspace topology induced by the Euclidean topology on $\mathbb{R}^n$.

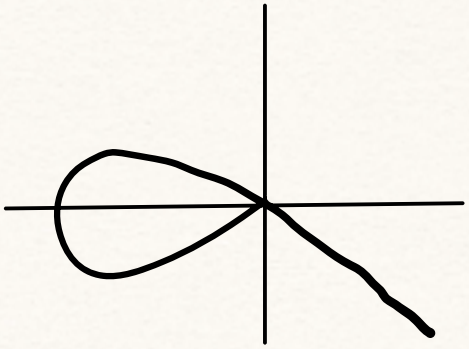
Definition: The **relative boundary** of X inside $\overline{X}^{Zar}$, denoted by ∂X, is the boundary of X in the Euclidean subspace topology on $\overline{X}^{Zar}$.

Definition: If $\dim\left(\bigcup_{p \in \partial X} Var_X(p)\right) = n$, then ∂X is **exposed** w.r.t. Euclidean distance minimization to X.

Relative boundary

Consider the curve C defined by $y^2 - x^2(x+1) = 0$:



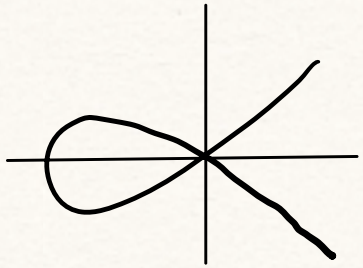


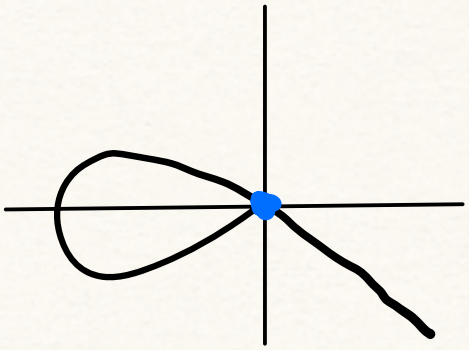
What is ∂X?

$X = \{(x,y) \in C \mid x \leq 0 \text{ or } y \leq 0\}$

Relative boundary

Consider the curve C defined by $y^2 - x^2(x+1) = 0$:





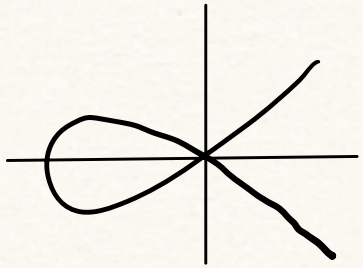
$\partial X = \{(0,0)\}$

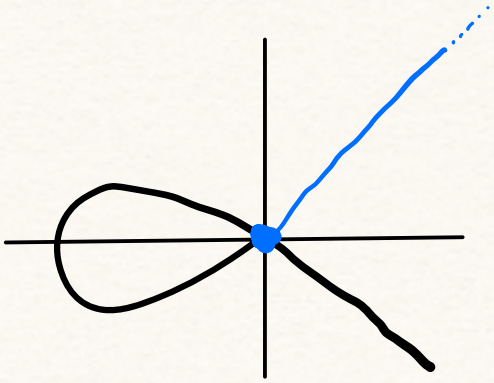
What is $\mathrm{Vor}_X((0,0))$?

$X = \{(x,y) \in C \mid x \in 0 \text{ or } y \in 0\}$

Relative boundary

Consider the curve C defined by $y^2 - x^2(x+1) = 0$:





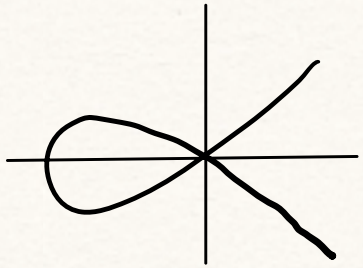
$\partial X = \{(0,0)\}$

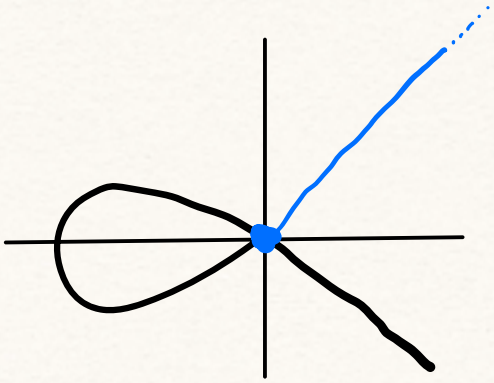
$\mathrm{Vor}_X((0,0)) = \{(x,x) \mid x \geq 0\}$

$X = \{(x,y) \in C \mid x \leq 0 \text{ or } y \leq 0\}$

Relative boundary

Consider the curve C defined by $y^2 - x^2(x+1) = 0$:





$\partial X = \{(0,0)\}$

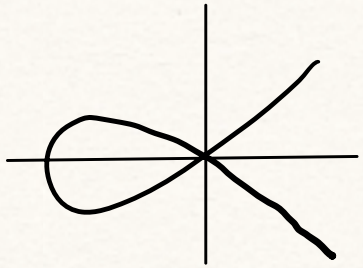
$\mathrm{Vor}_X((0,0)) = \{(x,x) \mid x \geq 0\}$

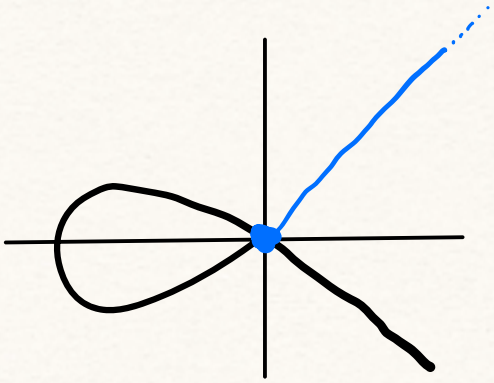
$X = \{(x,y) \in C \mid x \leq 0 \text{ or } y \leq 0\}$

If ∂X is not exposed:
For almost all $u \in \mathbb{R}^n$,
$\operatorname{argmin}_{x \in X} \|x-u\|^2 \subseteq X \setminus \partial X$.

Relative boundary

Consider the curve C defined by $y^2 - x^2(x+1) = 0$:



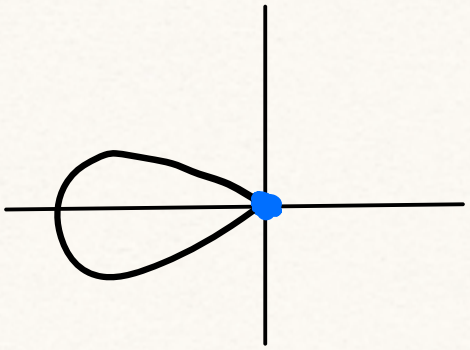


$\partial X = \{(0,0)\}$

$\mathrm{Vor}_X((0,0)) = \{(x,x) \mid x \geq 0\}$

$X = \{(x,y) \in C \mid x \leq 0 \text{ or } y \leq 0\}$

If ∂X is not exposed:
For almost all $u \in \mathbb{R}^n$,
$\operatorname{argmin}_{x \in X} \|x-u\|^2 \subseteq X \setminus \partial X$.



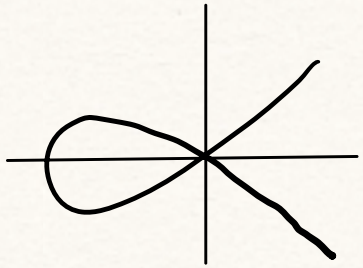
$\partial X = \{(0,0)\}$

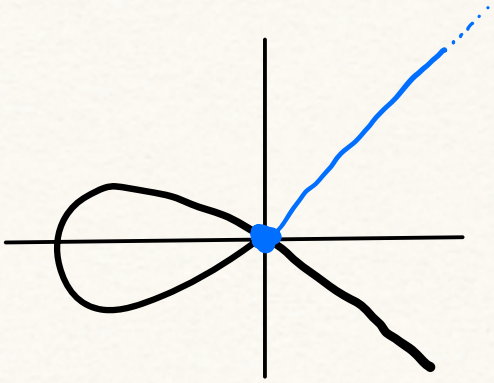
What is $\mathrm{Vor}_X((0,0))$?

$X = \{(x,y) \in C \mid x \leq 0\}$

Relative boundary

Consider the curve C defined by $y^2 - x^2(x+1) = 0$:



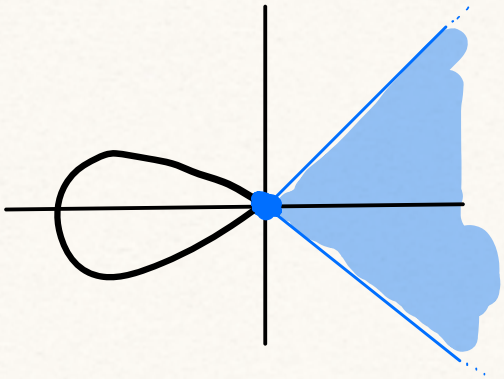


$\partial X = \{(0,0)\}$

$\mathrm{Vor}_X((0,0)) = \{(x,x) \mid x \geq 0\}$

$X = \{(x,y) \in C \mid x \leq 0 \text{ or } y \leq 0\}$

If ∂X is not exposed:
For almost all $u \in \mathbb{R}^n$,
$\operatorname{argmin}_{x \in X} \|x-u\|^2 \subseteq X \setminus \partial X$.



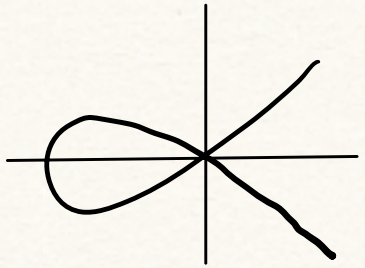
$\partial X = \{(0,0)\}$

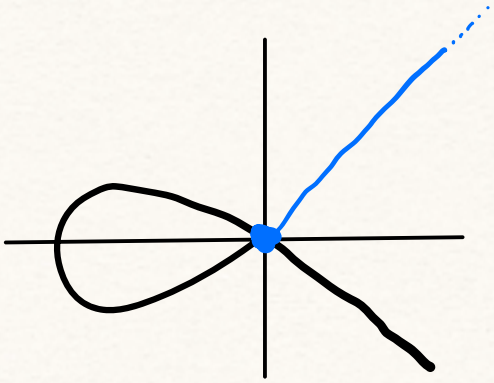
$\mathrm{Vor}_X((0,0)) = \{(x,y) \mid x \geq |y|\}$

$X = \{(x,y) \in C \mid x \leq 0\}$

Relative boundary

Consider the curve C defined by $y^2 - x^2(x+1) = 0$:



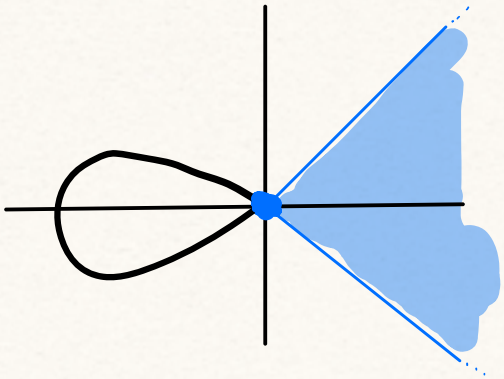


$\partial X = \{(0,0)\}$

$\mathrm{Vor}_X((0,0)) = \{(x,x) \mid x \geq 0\}$

$X = \{(x,y) \in C \mid x \leq 0 \text{ or } y \leq 0\}$

If ∂X is not exposed:
For almost all $u \in \mathbb{R}^n$,
$\operatorname{argmin}_{x \in X} \|x-u\|^2 \subseteq X \setminus \partial X$.



$\partial X = \{(0,0)\}$

$\mathrm{Vor}_X((0,0)) = \{(x,y) \mid x \geq |y|\}$

$X = \{(x,y) \in C \mid x \leq 0\}$

If ∂X is exposed:
For all data distributions on $\mathbb{R}^n$ with full support, $\operatorname{argmin}_{x \in X} \|x-u\|^2$ contains a point on ∂X with positive probability on the data $u \in \mathbb{R}^n$.

Boundary Stratification

Assume: noisy data $u \in \mathbb{R}^n$, $X \subseteq \mathbb{R}^n$ semialgebraic

Goal: find $\operatorname{argmin}_{x \in X} \|x-u\|^2$

If ∂X is not exposed: enough to find critical points of $\min_{x \in \overline{X}^{Zar}} \|x-u\|^2$.

If ∂X is exposed:

need to find critical points of $\min_{x \in \overline{X}^{Zar}} \|x-u\|^2$ and solve $\operatorname{argmin}_{x \in \partial X} \|x-u\|^2$.

Boundary Stratification

Assume: noisy data $u \in \mathbb{R}^n$, $X \subseteq \mathbb{R}^n$ semialgebraic

Goal: find $\operatorname{argmin}_{x \in X} \|x-u\|^2$

If ∂X is not exposed: enough to find critical points of $\min_{x \in \overline{X}^{Zar}} \|x-u\|^2$.

If ∂X is exposed:

need to find critical points of $\min_{x \in \overline{X}^{Zar}} \|x-u\|^2$ and solve $\operatorname{argmin}_{x \in \partial X} \|x-u\|^2$.

↑ ∂X is semialgebraic, so we can iterate this process ☺

Boundary Stratification

Assume: noisy data $u \in \mathbb{R}^n$, $X \subseteq \mathbb{R}^n$ semialgebraic

Goal: find $\operatorname{argmin}_{x \in X} \|x-u\|^2$

If ∂X is not exposed: enough to find critical points of $\min_{x \in \overline{X}^{Zar}} \|x-u\|^2$.

If ∂X is exposed:

need to find critical points of $\min_{x \in \overline{X}^{Zar}} \|x-u\|^2$ and solve $\operatorname{argmin}_{x \in \partial X} \|x-u\|^2$.

↑ ∂X is semialgebraic, so we can iterate this process :)

Definition: The Zariski closure of ∂X inside $\mathbb{R}^n$ is called the algebraic boundary of X, denoted by $\partial_{alg} X$.

Example: A Neuromanifold

3 Jacobian check

Week 3

Exercise 5. *Consider the following* shallow cubic neural network*:*

$$\begin{bmatrix} e & f \end{bmatrix} \sigma\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}\right), \quad \text{where } \sigma\left(\begin{bmatrix} u \\ v \end{bmatrix}\right) := \begin{bmatrix} u^3 \\ v^3 \end{bmatrix},$$

x, y are inputs to the network, and $a, b, \ldots, f$ are network parameters. Note that the left-hand side above equals

$$(a^3e + c^3f)x^3 + 3(a^2be + c^2df)x^2y + 3(ab^2e + cd^2f)xy^2 + (b^3e + d^3f)y^3. \quad =: F$$

Show that, when varying the network parameters, the network parametrizes a 4-dimensional family of functions in x and y.

$\mathbb{R}^6 \longrightarrow \mathbb{R}[x,y]_3 \cong \mathbb{R}^4$

$(a,b,c,d,e,f) \mapsto F$

Example: A Neuromanifold

3 Jacobian check

Week 3

Exercise 5. *Consider the following* shallow cubic neural network*:*

$$\begin{bmatrix} e & f \end{bmatrix} \sigma\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}\right), \quad \text{where } \sigma\left(\begin{bmatrix} u \\ v \end{bmatrix}\right) := \begin{bmatrix} u^3 \\ v^3 \end{bmatrix},$$

x, y are inputs to the network, and $a, b, \ldots, f$ are network parameters. Note that the left-hand side above equals

$$(a^3e + c^3f)x^3 + 3(a^2be + c^2df)x^2y + 3(ab^2e + cd^2f)xy^2 + (b^3e + d^3f)y^3. \quad =: F$$

Show that, when varying the network parameters, the network parametrizes a 4-dimensional family of functions in x and y.

$\mathbb{R}^6 \longrightarrow \mathbb{R}[x,y]_3 \cong \mathbb{R}^4$

$(a,b,c,d,e,f) \mapsto F$

The image of this map is called the neuromanifold $\mathcal{M}$ of the network.

The exercise shows that $\overline{\mathcal{M}}^{\text{Zar}} = \mathbb{R}[x,y]_3$.

Example: A Neuromanifold

3 Jacobian check

Week 3

Exercise 5. *Consider the following* shallow cubic neural network*:*

$$\begin{bmatrix} e & f \end{bmatrix} \sigma\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}\right), \quad \text{where } \sigma\left(\begin{bmatrix} u \\ v \end{bmatrix}\right) := \begin{bmatrix} u^3 \\ v^3 \end{bmatrix},$$

x, y are inputs to the network, and $a, b, \ldots, f$ are network parameters. Note that the left-hand side above equals

$$(a^3e + c^3f)x^3 + 3(a^2be + c^2df)x^2y + 3(ab^2e + cd^2f)xy^2 + (b^3e + d^3f)y^3. \quad =: F$$

Show that, when varying the network parameters, the network parametrizes a 4-dimensional family of functions in x and y.

$\mathbb{R}^6 \longrightarrow \mathbb{R}[x,y]_3 \cong \mathbb{R}^4$

$(a,b,c,d,e,f) \mapsto F$

The image of this map is called the neuromanifold $\mathcal{M}$ of the network.

What is the actual $\mathcal{M}$ (without taking closures)?

The exercise shows that $\overline{\mathcal{M}}^{\text{Zar}} = \mathbb{R}[x,y]_3$.

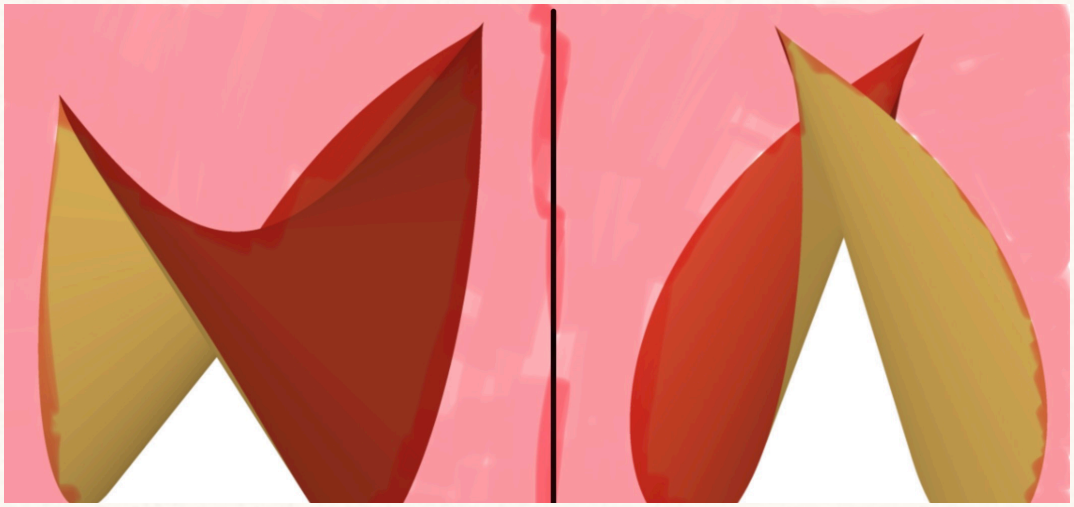
Example: A Newsomanifold

$$\Delta := \{F \in \mathbb{R}[x,y]_3 \mid F \text{ has a double root}\}$$

$$= \{Ax^3 + Bx^2y + Cxy^2 + Dy^3 \mid \underbrace{B^2C^2 - 4AC^3 - 4B^3D - 27A^2D^2 + 18ABCD}_{\delta} = 0\}$$

$$\mathrm{Sing}(\Delta) = \{F \in \mathbb{R}[x,y]_3 \mid F \text{ has a triple root}\} = \{\ell^3 \mid \ell \in \mathbb{R}[x,y]_1\}$$

$$\mathcal{M} = \{F \in \mathbb{R}[x,y]_3 \mid \delta(F) < 0\} \cup \mathrm{Sing}(\Delta)$$



Example: A Newsomanifold

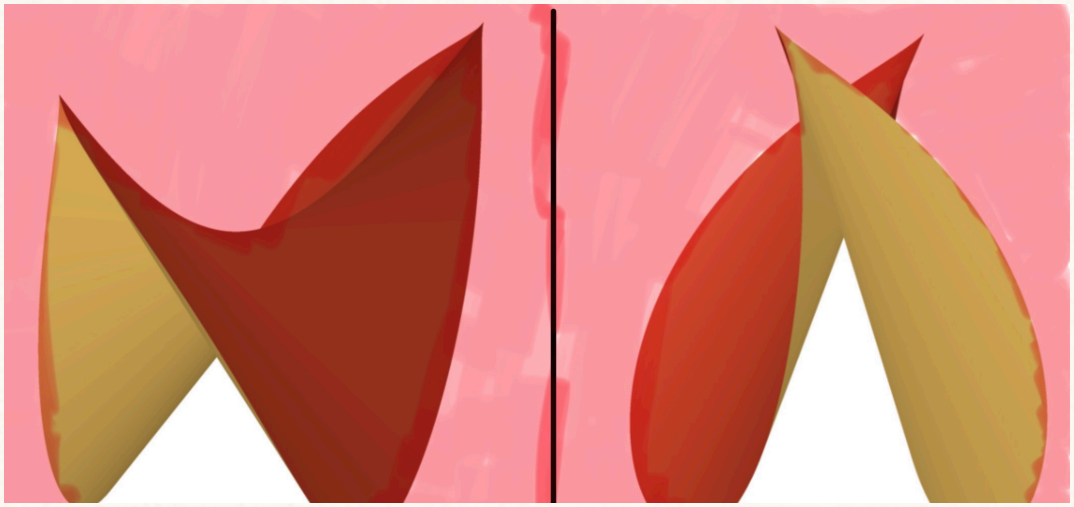
$$\Delta := \{F \in \mathbb{R}[x,y]_3 \mid F \text{ has a double root}\}$$

$$= \{Ax^3 + Bx^2y + Cxy^2 + Dy^3 \mid \underbrace{B^2C^2 - 4AC^3 - 4B^3D - 27A^2D^2 + 18ABCD}_{\delta} = 0\}$$

$$\mathrm{Sing}(\Delta) = \{F \in \mathbb{R}[x,y]_3 \mid F \text{ has a triple root}\} = \{\ell^3 \mid \ell \in \mathbb{R}[x,y]_1\}$$

$$\mathcal{M} = \{F \in \mathbb{R}[x,y]_3 \mid \delta(F) < 0\} \cup \mathrm{Sing}(\Delta)$$

$$\partial\mathcal{M} = ?$$



Example: A Newsomanifold

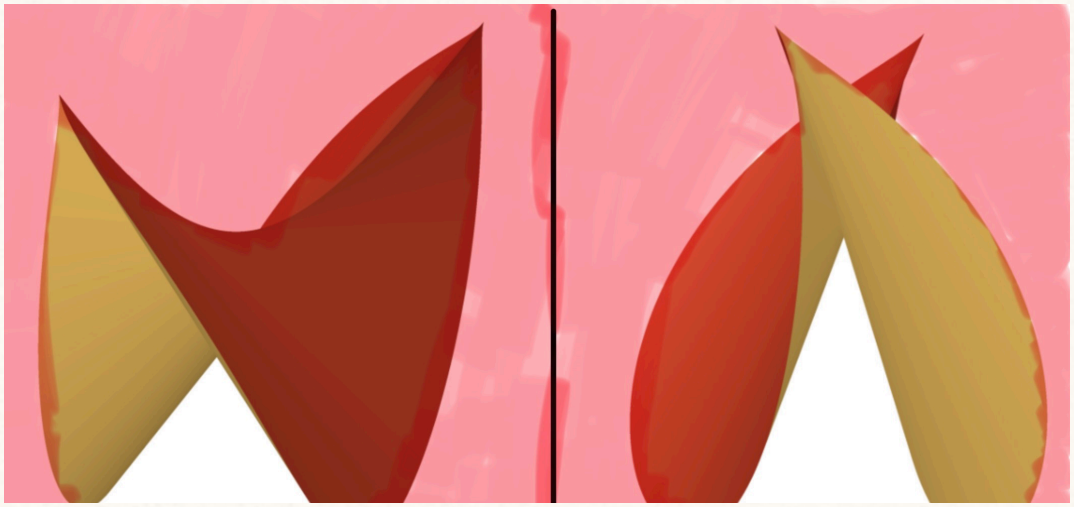
$$\Delta := \{F \in \mathbb{R}[x,y]_3 \mid F \text{ has a double root}\}$$

$$= \{Ax^3 + Bx^2y + Cxy^2 + Dy^3 \mid \underbrace{B^2C^2 - 4AC^3 - 4B^3D - 27A^2D^2 + 18ABCD}_{\delta} = 0\}$$

$$\mathrm{Sing}(\Delta) = \{F \in \mathbb{R}[x,y]_3 \mid F \text{ has a triple root}\} = \{\ell^3 \mid \ell \in \mathbb{R}[x,y]_1\}$$

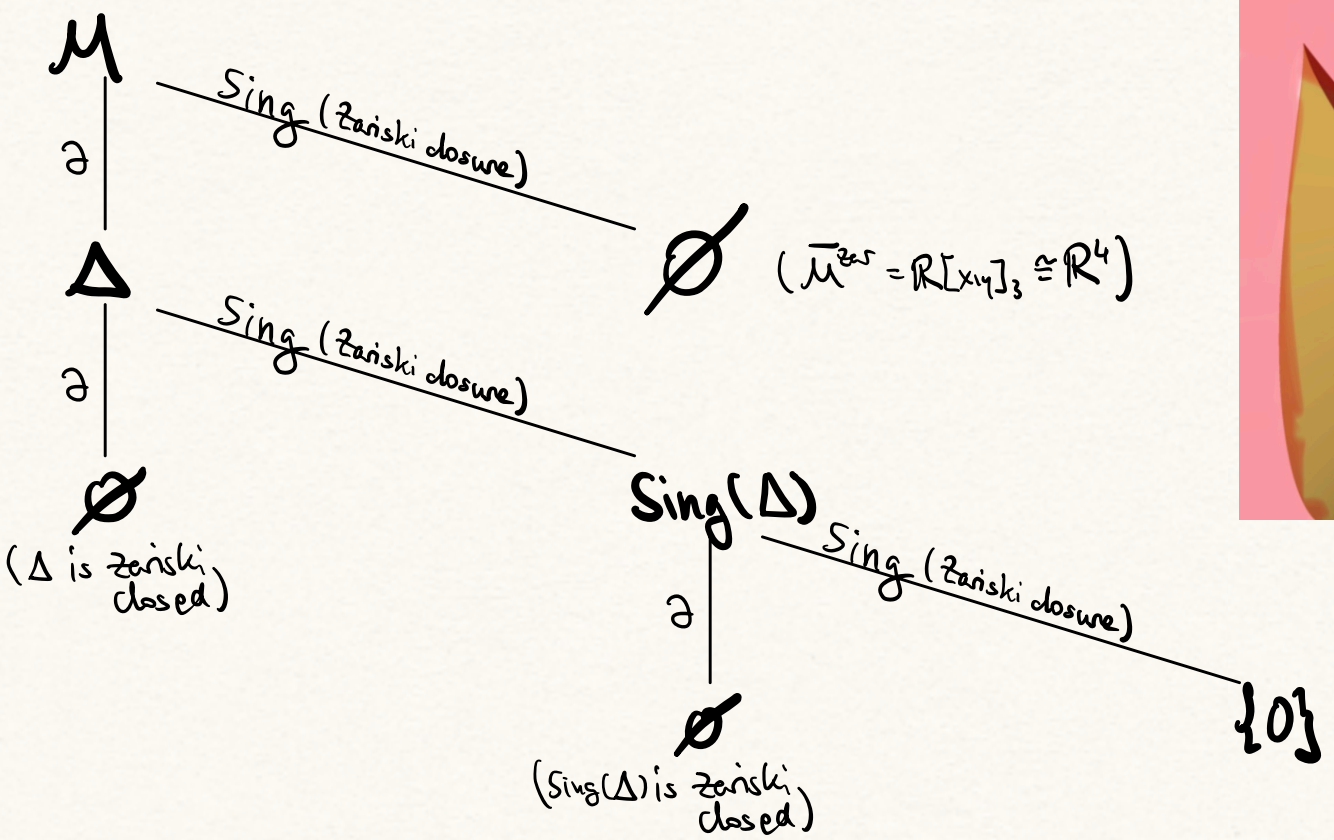
$$M = \{F \in \mathbb{R}[x,y]_3 \mid \delta(F) < 0\} \cup \mathrm{Sing}(\Delta)$$

$$\partial M = \Delta$$

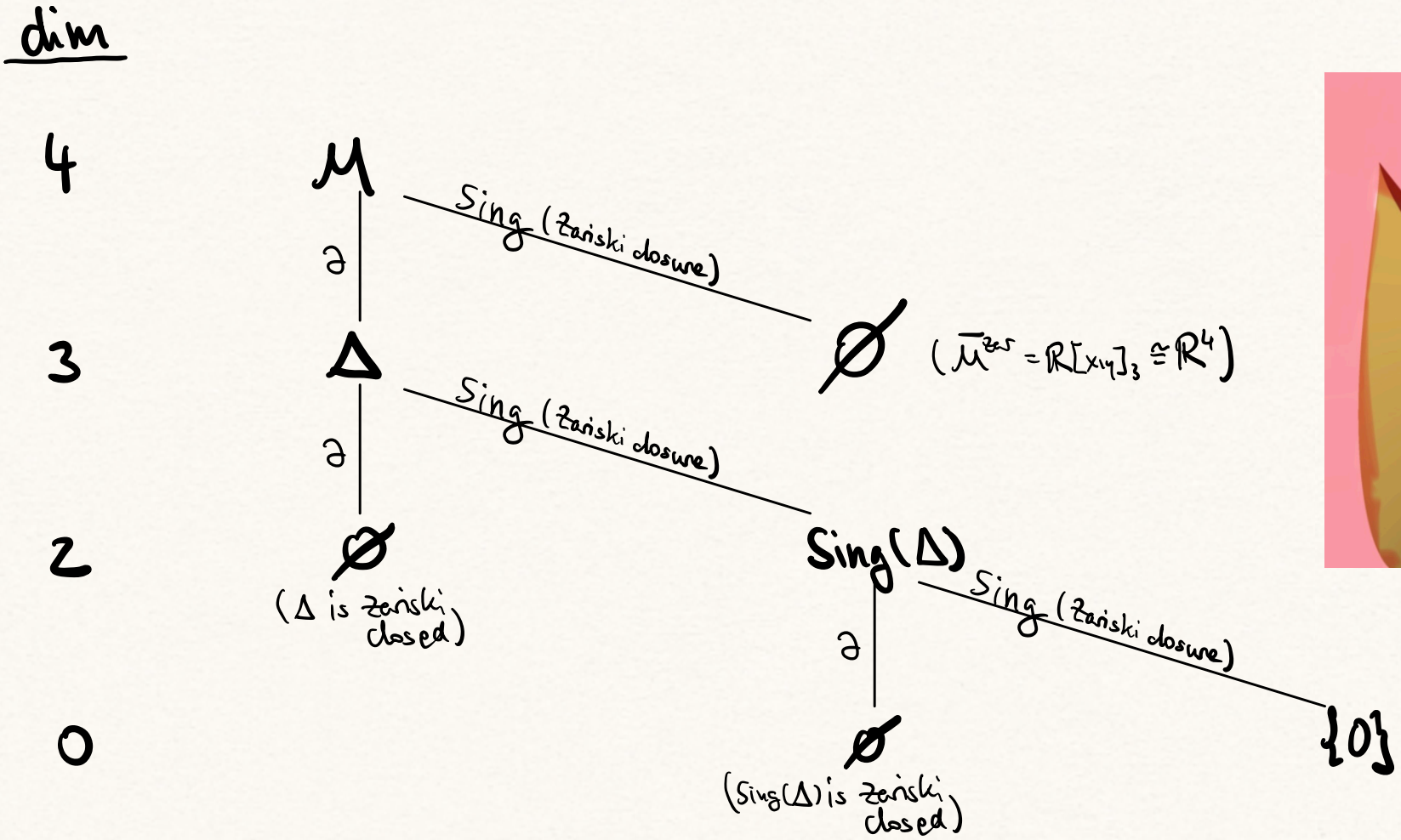


Example: A Neuromanifold → Stratification





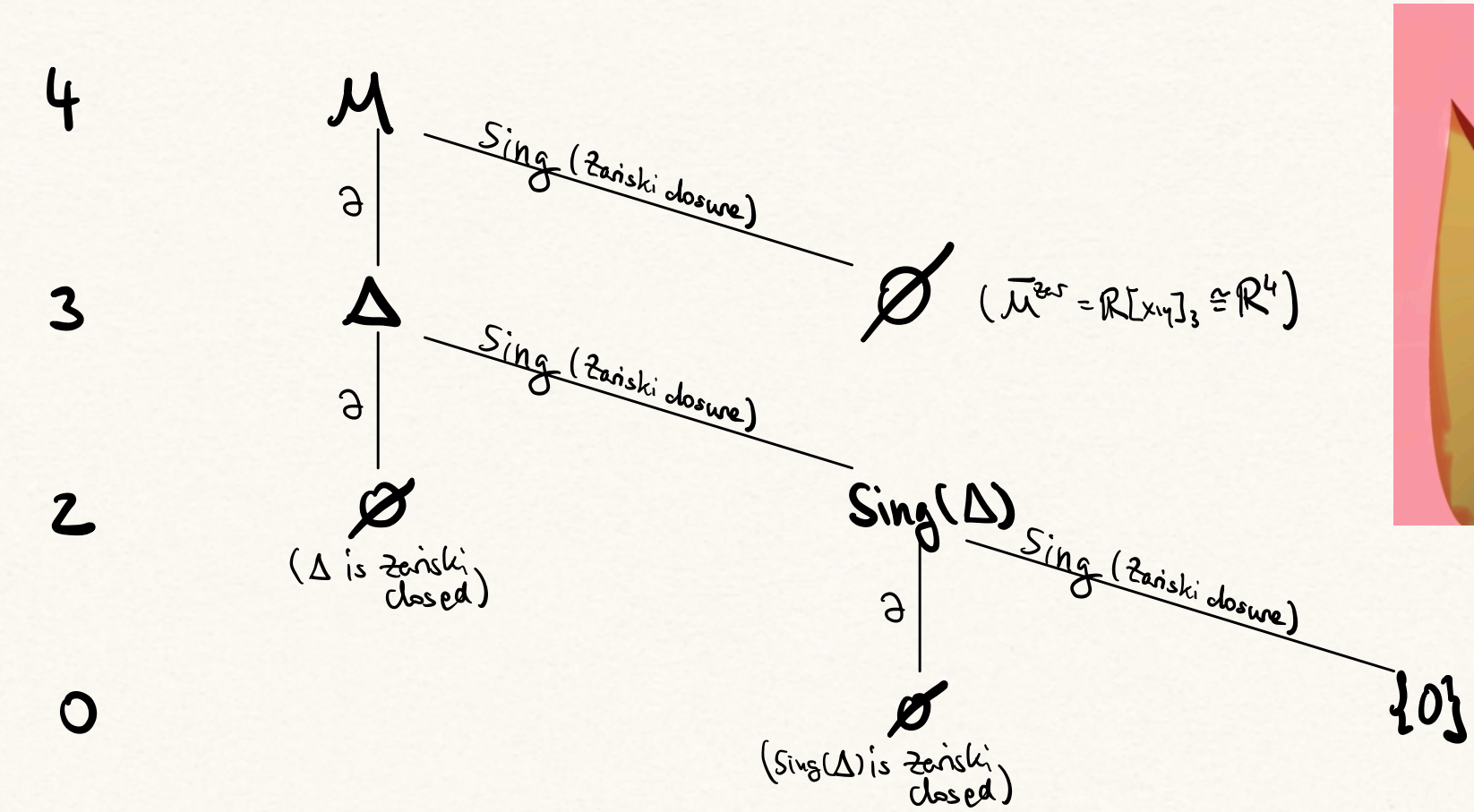
Example: A Newsomanifold → Stratification





- $\forall F \in \mathrm{Sing}(\Delta) : \mathrm{Vor}_M(F) = \{F\}$ ⇒ Sing(Δ) & {0} are not exposed

Example: A Newsomanifold → Stratification





- $\forall F \in \mathrm{Sing}(\Delta) : \mathrm{Vor}_M(F) = \{F\}$ $\Rightarrow$ Sing(Δ) & {0} are **not exposed**
- $\bigcup_{F \in \Delta} \mathrm{Vor}_M(F) = \{ G \in \mathbb{R}[x_{1:4}]_3 \mid \delta(G) \geq 0 \}$ $\Rightarrow$ Δ is **exposed**

Example: A Neuromanifold → Euclidean distance minimization



When solving $\inf_{F \in \mathcal{M}} \|F - G\|^2$ for noisy data G,

we need to consider $\text{int}(\mathcal{M})$ & (the smooth locus of) Δ:

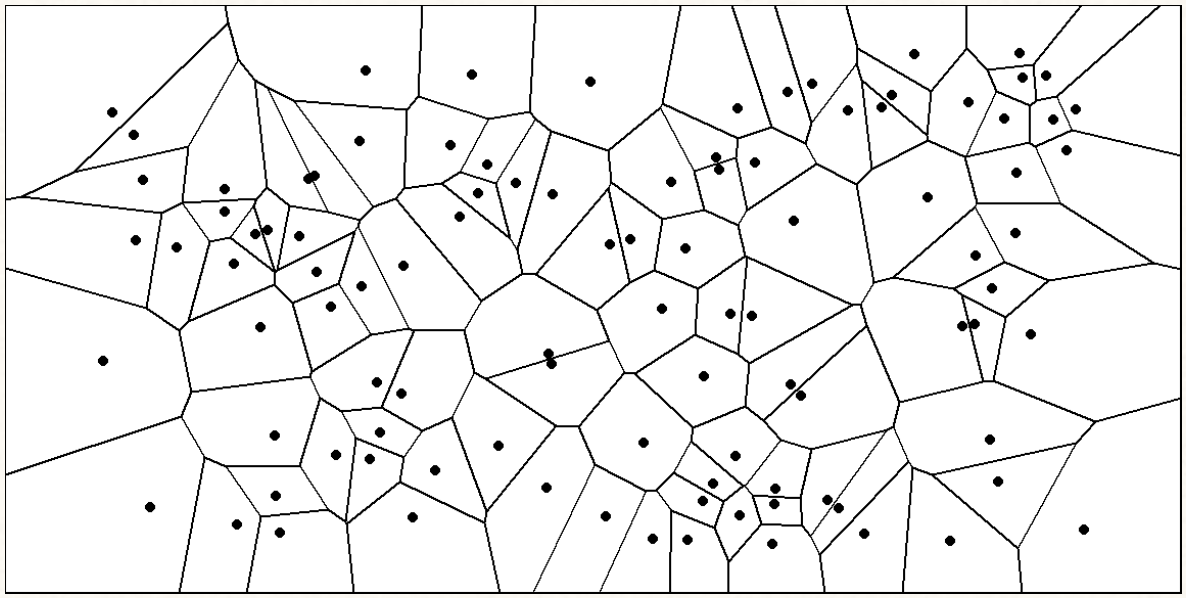
1) If $\delta(G) < 0$: $G \in \mathcal{M}$ is minimizes.

2) If $\delta(G) > 0$: infimum lies on smooth locus of Δ

↑

Exercise: $EDD(\Delta) = 7$

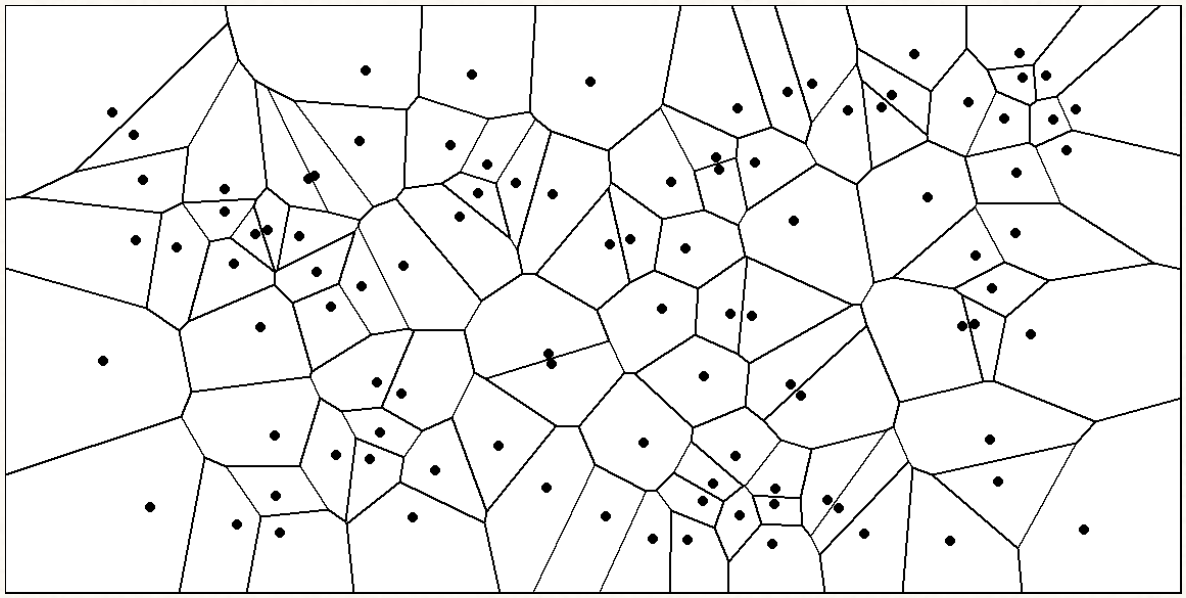
Medial Axis



The Voronoi diagram of a finite point set X is formed by the boundaries of its Voronoi cells.

What is the analog for continuous sets X?

Medial Axis



The Voronoi diagram of a finite point set X is formed by the boundaries of its Voronoi cells.

For $X \subseteq \mathbb{R}^n$ semialgebraic:

$$\bigcup_{x \in X} \partial \, \mathrm{Vor}_X(x) =$$

$$\{u \in \mathbb{R}^n : \log \min_{x \in X} \|x-u\|^2 | > 1\}$$

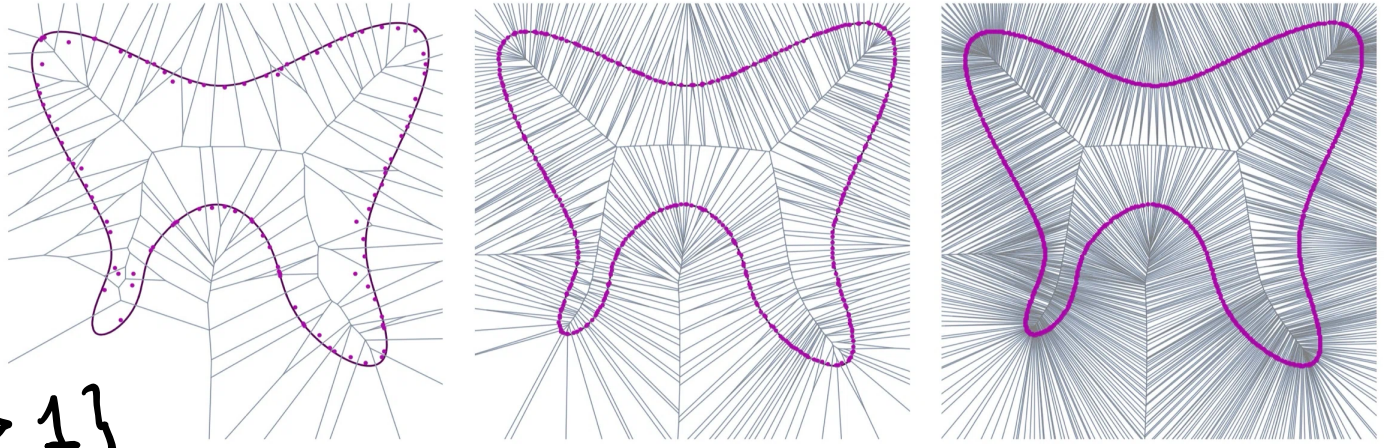


Fig. 8.1: We see the quartic curve $X = \{x^4 - x^2y^2 + y^4 - 4x^2 - 2y^2 - x - 4y + 1 = 0\}$, called *butterfly*

This is known as the medial axis of X.