Data Fitting (cont'd)

Q: How to measure goodness of (x,t)?

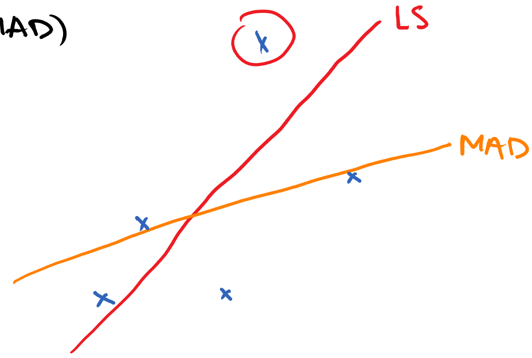
Approach 1: Least-squares approach

$$\min_{\substack{x\in\mathbb{R}^n\\ t\in\mathbb{R}}} \sum_{i=1}^{m} (\underbrace{a_i^T x + t}_{\substack{f(a_i)\\ \text{predicted}\\ \text{output of } a_i}} - \underbrace{b_i}_{\substack{\text{actual}\\ \text{output}}})^2$$

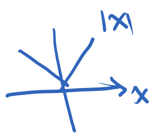
Compute the gradient, set it to zero to get the optimal solution (exercise)

Approach 2: Minimum absolute deviation (MAD)

$$\min_{\substack{x\in\mathbb{R}^n\\ t\in\mathbb{R}}} \sum_{i=1}^{m} |a_i^T x + t - b_i|$$



Note: $x \mapsto |x|$ is not differentiable at $x=0$



Consider

$$\min \sum_{i=1}^{m} z_i$$

$$\text{s.t.} \quad z_i = |a_i^T x + t - b_i|, \quad i=1,\cdots,m$$

can change to $\geq$ (convince yourself)

Note:

$$z_i \geq |a_i^T x + t - b_i| \iff z_i \geq a_i^T x + t - b_i \geq -z_i$$

Hence, we obtain

$$\min \sum_{i=1}^{m} z_i$$

$$\text{s.t.} \quad \left.\begin{array}{l} z_i \geq a_i^T x + t - b_i \\ z_i \geq -(a_i^T x + t - b_i) \end{array}\right\} i=1,\cdots,m$$

$$x\in\mathbb{R}^n,\ z\in\mathbb{R}^m,\ t\in\mathbb{R}$$

This is an LP.

Convexity

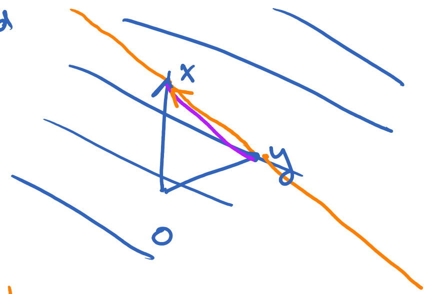
Let $S \subseteq \mathbb{R}^n$ be arbitrary.

Definitions:

① S is a linear subspace if $\forall x, y \in S$; $\alpha, \beta \in \mathbb{R}$,

$$\alpha x + \beta y \in S$$

plane spanned by x and y



② S is affine if $\forall x, y \in S$; $\alpha \in \mathbb{R}$,

$$\underbrace{\alpha x + (1-\alpha) y}_{= y + \alpha(x-y)} \in S$$

line through x and y

③ S is convex if $\forall x, y \in S$, $\alpha \in [0,1]$,

$$\alpha x + (1-\alpha) y \in S$$

line segment between x and y

Example: Consider $S = \{x\}$. Is this linear? affine? convex?

1) Linear if $x = 0$, o/w not linear.

2) Affine always 3) convex always

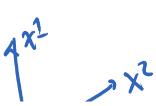
④ Given $x^1, \dots, x^k \in \mathbb{R}^n$, we say $y = \sum_{i=1}^{k} \alpha_i x^i$ is

a) linear combination of $x^1, \dots, x^k$ if $\alpha_1, \dots, \alpha_k \in \mathbb{R}$;

b) affine combination of $x^1, \dots, x^k$ if $\alpha_1, \dots, \alpha_k \in \mathbb{R}$, $\sum_{i=1}^{k} \alpha_i = 1$;

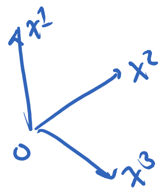
c) convex combination of $x^1, \dots, x^k$ if $\alpha_1, \dots, \alpha_k \geq 0$, $\sum_{i=1}^{k} \alpha_i = 1$.

e.g.



linear: space spanned x^1, x^2, x^3

e.g.



linear : space spanned x^1, x^2, x^3

affine : plane through x^1, x^2, x^3

Convex : triangle with corners x^1, x^2, x^3

How does an affine set look like?

<u>Proposition</u>: The following are equivalent:

1) S is affine

2) Any affine combination of a finite number of points in S belongs to S

3) S can be written as $S = \underbrace{\{x\}}_{\text{shift}} + \underbrace{V}_{\text{linear subspace}} \triangleq \{x+v : v \in V\}$

for some $x \in \mathbb{R}^n$ and linear subspace V.

<u>e.g.</u> $\mathbb{R}^2$: 3 types of linear subspace

1) 0-dim : $V = \{0\}$

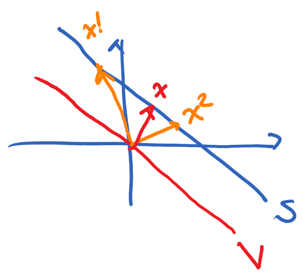
2) 1-dim : Span by a single vector $x \neq 0$

$$V = \{\alpha x : \alpha \in \mathbb{R}\}$$

3) 2-dim : span by two linearly independent vectors x, y

$$V = \{\alpha x + \beta y : \alpha, \beta \in \mathbb{R}\} = \mathbb{R}^2.$$

e.g.



<u>Note</u>: Here, the linear subspace V is unique but the shift x is not

<u>Example</u>: Let $S = \{x \in \mathbb{R}^n : Ax = b\}$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $m \leq n$

<u>Claim</u>: S is affine.

<u>Proof</u>: Take any $x, y \in S$; $\alpha \in \mathbb{R}$.

<u>Want</u>: $z = \alpha x + (1-\alpha) y \in S$.

Want: $z = \alpha x + (1-\alpha) y \in S$.

① $x \in S \Leftrightarrow Ax = b$; ② $y \in S \Leftrightarrow Ay = b$

Compute

$$Az = A(\alpha x + (1-\alpha) y) = \alpha \underbrace{Ax}_{\substack{= b \\ \text{by } ①}} + (1-\alpha) \underbrace{Ay}_{\substack{= b \\ \text{by } ②}} = b$$

$\Leftrightarrow z \in S$.

How about convex sets?

Proposition:

S is convex iff any convex combination of a finite number of points in S belongs to S.