Lecture 3

Today: mostly examples

Last time: ℓ_1, ℓ_∞ norm minimization – equivalent to LP

$x \in \mathbb{R}^n$, $\quad \min_x \|x\|_\infty$ $\quad$ ℓ_∞ infinity norm

$$\|x\|_\infty = \max\{|x_1|, |x_2|, \dots, |x_n|\}$$

$$= \max\{x_1, -x_1, x_2, -x_2, \dots, x_n, -x_n\}$$

$$\begin{aligned} &\min t \\ &\text{st} \quad x_i \le t \qquad i = 1, \dots, n \\ &\qquad -x_i \le t \end{aligned}$$

ℓ_1-norm $\quad \min_x \|x\|_1$

$x \in \mathbb{R}^n$

$$\|x\|_1 = \sum_{i=1}^n |x_i| = \sum_{i=1}^n \max\{x_i, -x_i\}$$

$$\min_{x, t_1, \dots, t_n} t_1 + t_2 + \dots + t_n$$

$$\text{s.t.}\quad \left.\begin{aligned} x_i &\le t_i \\ -x_i &\le t_i \end{aligned}\right\} \; -t_i \le x_i \le t_i$$

$$i = 1, \dots, n.$$

$$\min \; 1^T t \qquad t = \begin{pmatrix} t_1 \\ \vdots \\ t_n \end{pmatrix}$$

$$\text{s.t.} \quad -t \le x \le t$$

- other norms: use approximation

$$A \|x\|_p \le \|x\|_q \le B \|x\|_p$$

<u>Example</u>

$$\|x\|_2 \le \|x\|_1 \le \sqrt{n}\, \|x\|_2$$

$$\begin{array}{l} \min \ \|x\|_2 \\ \text{s.t.} \\ \text{some constraints} \\ \quad Ax \le b \end{array}$$

Optimal value p^*
at x_0^*

$$\begin{array}{l} \min \ \|x\|_1 \\ \text{s.t.} \\ \text{same constraints} \\ \quad Ax \le b \end{array}$$

Optimal value q^*
at x_1^*

How far is q^* from p^*?

"Difference" between $\| \ \|_1$ and $\| \ \|_2$

- small entries contribute more to $\| \ \|_1$ and less to $\| \ \|_2$

$$x = (0.1 \ , \ 10)$$

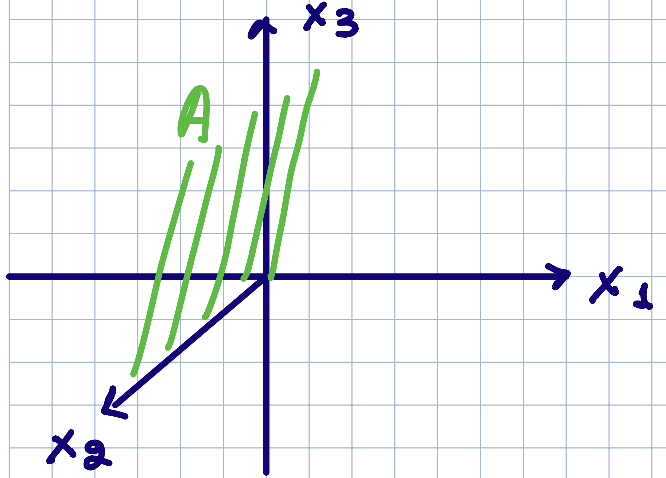
- large entries contribute more to $\| \ \|_2$

Error measurement

$\ell_2 \rightarrow$ lots of small errors

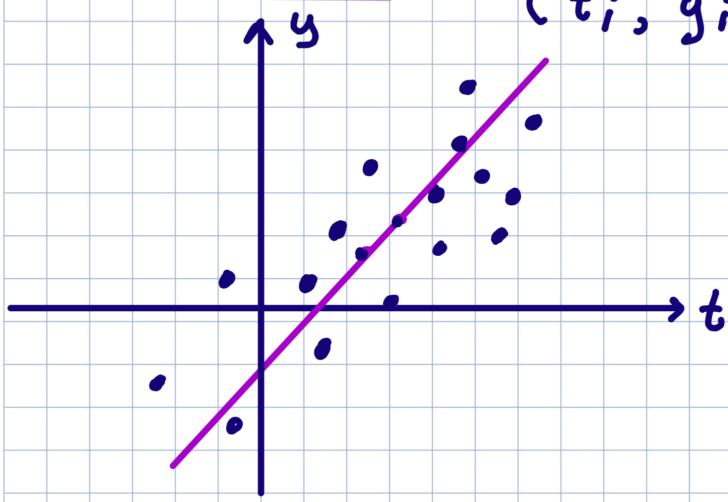
$\ell_1 \rightarrow$ most entries zero and a few larger ones

Example 1 : Assume we are given a vector $b \in \mathbb{R}^n$. We want to find the "closest" in the $\| \ \|_\infty$ sense, vector y, inside a subspace spanned by the columns of a matrix A.



$$A = \begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_k \\ | & | & & | \end{bmatrix}$$

Example 2



We are given m points $(t_i, y_i) \in \mathbb{R}^2$

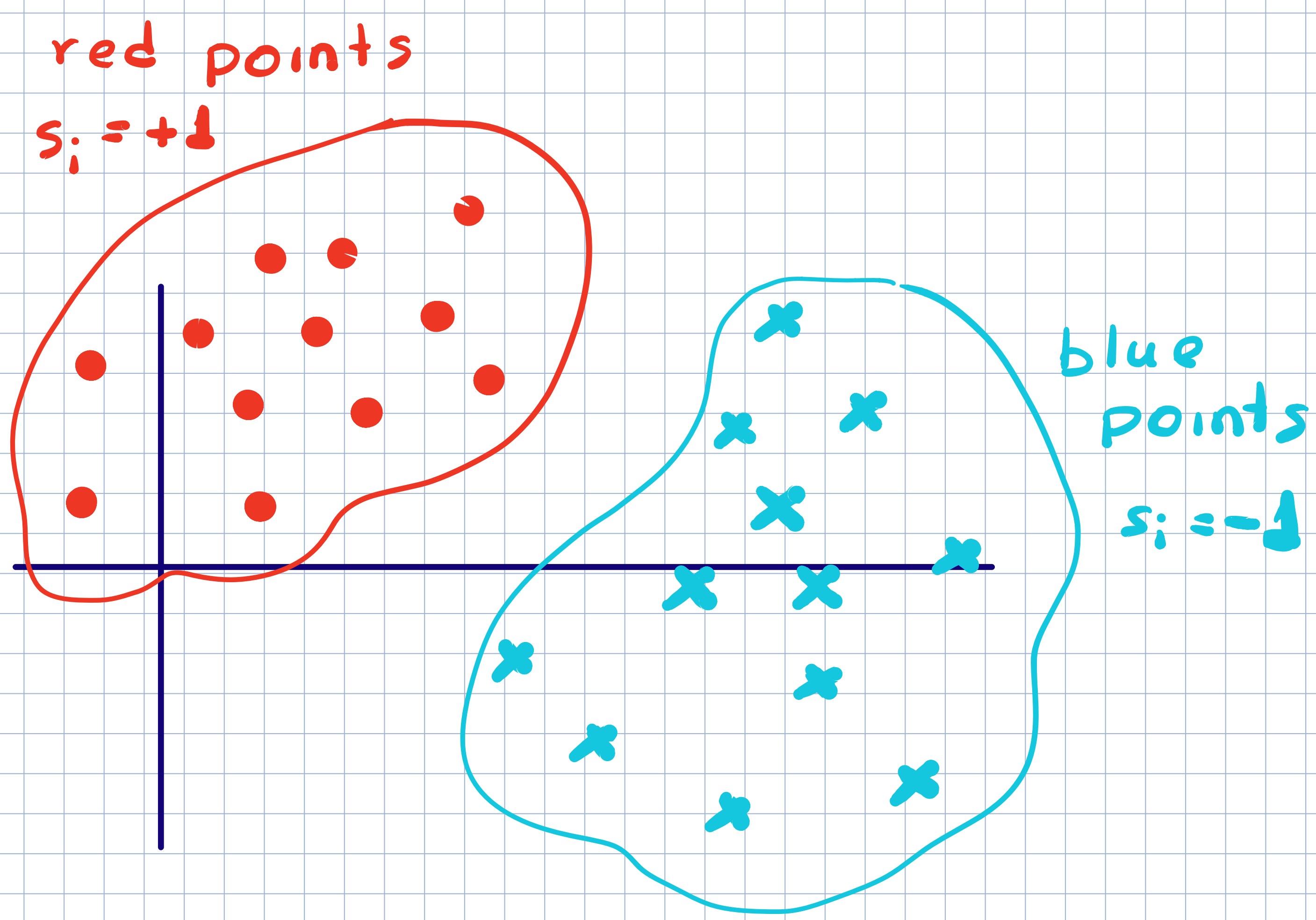
We want to find an affine function
$f(t) = at + b$
such that, for every t_i, $f(t_i)$ is close to y_i in ℓ_1 distance

Example 3 Linear Classification

Consider a set of points $x_1, x_2, \dots, x_m \in \mathbb{R}^n$

Each point comes with an associated label

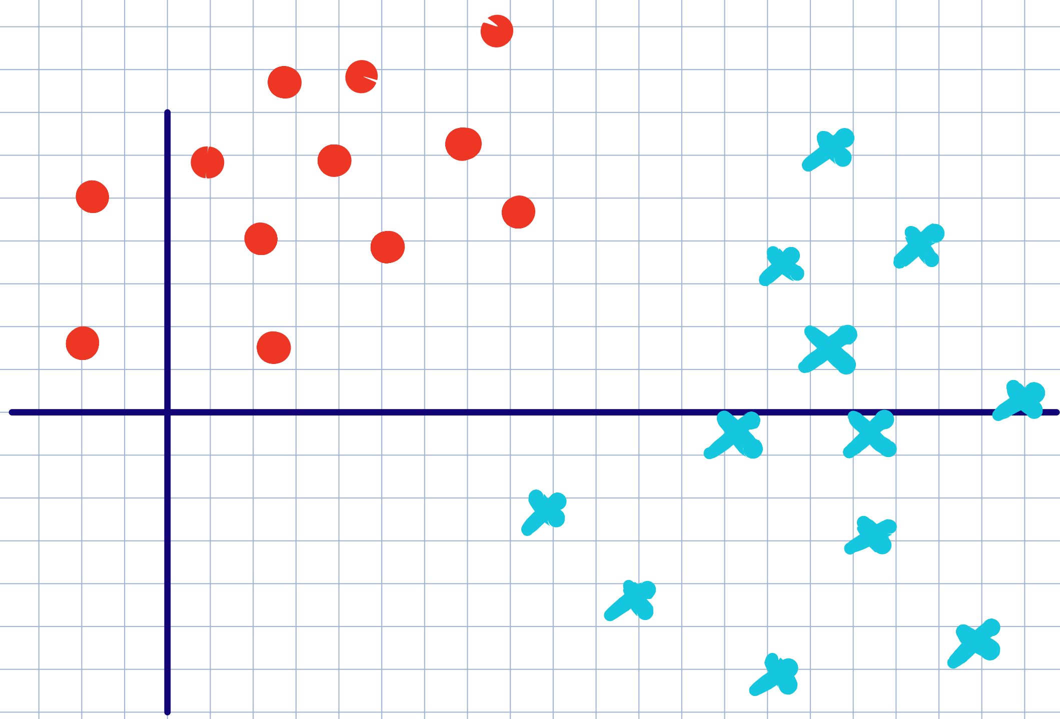
$s_i \in \{+1, -1\}$



Find a hyperplane $H = \{x \mid a^T x + b = 0\}$

that separates "as well as possible" the trainining points

If the training data are linearly separable (there exists a hyperplane that perfectly separates them) the objective value is zero.

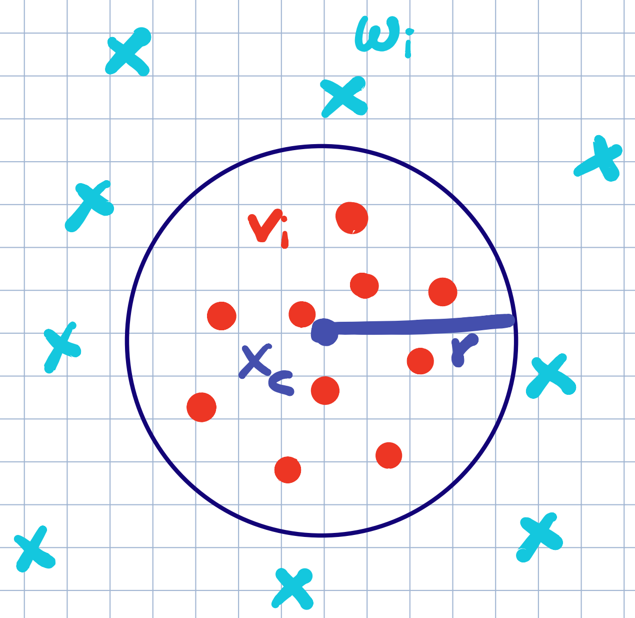


SVM (support vector machine):

maximize the "margin" around the separating hyperplanes

Example 4

Find a sphere that separates two sets of points v_i and ω_i



Sphere: all points within distance r from a center x_c.

$$S = \{ x \mid \| x - x_c \|_2 \le r \}$$