Approximate Solutions, Norms, and the Least-Squares Problem

Objectives

- state need for approximate solutions
- introduce norms
- define properties of norms
- introduce the least-squares problem

Solving Systems of Linear Equations is Important 2

$\underline{A}\,\underline{w} = \underline{d}$

- classification
- modeling

Exact solution: $\underline{d}$ must lie in the subspace spanned by the columns of $\underline{A}$

$\sum_i \underline{a}_i w_i = \underline{d}$

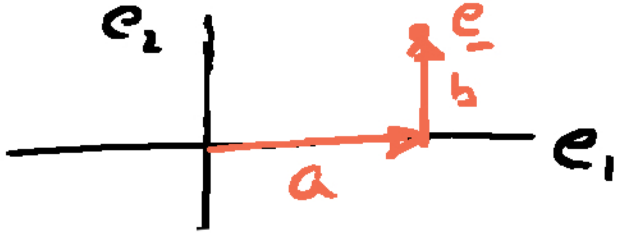
Rarely satisfied in real problems due to noise, model limitations, nonidealities, etc

Can we find $\underline{w}$ so $\underline{A}\,\underline{w} \approx \underline{d}$?

Error $\underline{e} = \underline{A}\,\underline{w} - \underline{d}$, want $\underline{e}$ small

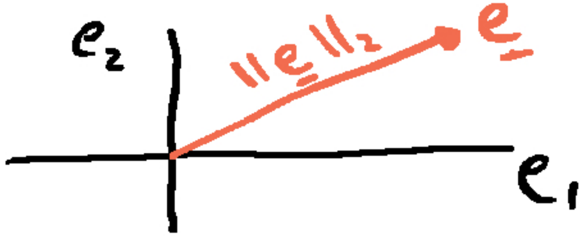
A vector norm measures the "size" of a vector.

- $\|\underline{e}\|_1 = \sum_i |e_i|$ $= a + b$



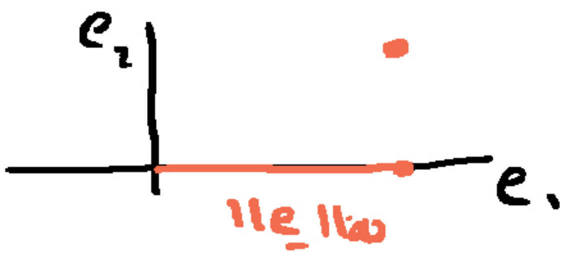
Taxi-cab norm ℓ_1

- $\|\underline{e}\|_2 = \left(\sum_i |e_i|^2 \right)^{1/2}$



Euclidean norm ℓ_2

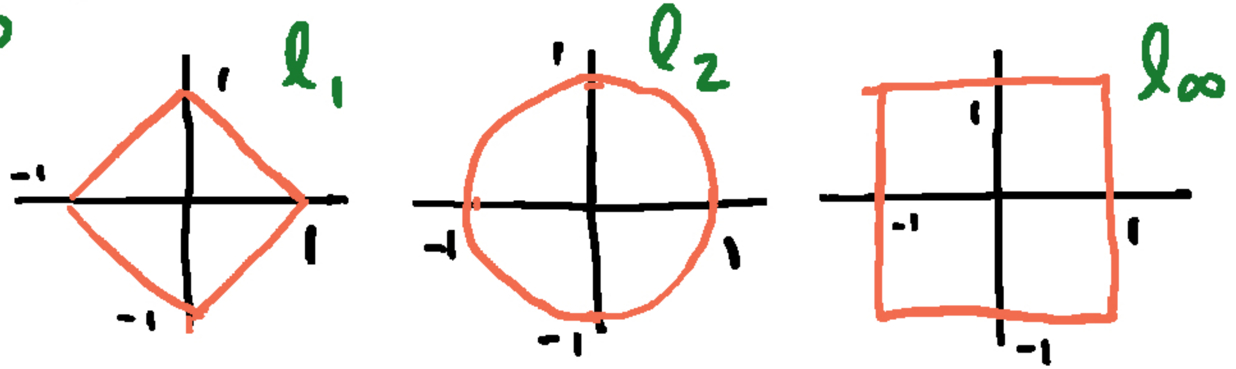
- $\|\underline{e}\|_\infty = \max_i |e_i|$



Sup or max norm ℓ_∞

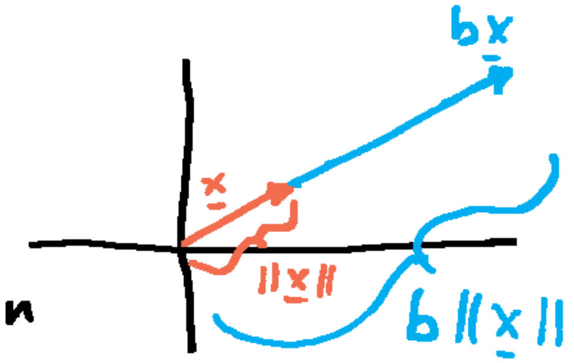
- $\|\underline{e}\|_q = \left(\sum_i |e_i|^q \right)^{1/q}$ ℓ_q norm

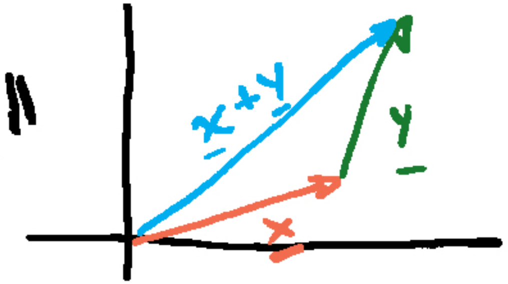
Unit ball: $\{\underline{x} : \|\underline{x}\| = 1\}$



A vector norm $\|\cdot\|$ maps from $\mathbb{R}^n \to \mathbb{R}$ and satisfies the following properties:

1) $\|\underline{x}\| \geq 0$ for all $\underline{x}$
2) $\|\underline{x}\| = 0$ if and only if $\underline{x} = \underline{0}$
3) $\|b\underline{x}\| = |b|\|\underline{x}\|$ for all $b \in \mathbb{R}$, $\underline{x} \in \mathbb{R}^n$
4) Triangle inequality $\|\underline{x} + \underline{y}\| \leq \|\underline{x}\| + \|\underline{y}\|$





Example: $\|\underline{x}\|_1 = \sum_i |x_i|$

1) $\|\underline{x}\|_1 \geq 0$

2) if $\underline{x} = \underline{0}$, then $\|\underline{x}\|_1 = 0$

if $\|\underline{x}\|_1 = \sum_i |x_i| = 0$, then $|x_i| = 0$

3) $\|b\underline{x}\|_1 = \sum_i |bx_i|$

$= |b| \sum_i |x_i|$

4) $\|\underline{x} + \underline{y}\|_1 = \sum_i |x_i + y_i|$

$\leq \sum_i |x_i| + |y_i| = \|\underline{x}\|_1 + \|\underline{y}\|_1$

Inclusion (ℓ_p norms): $1 \leq p \leq q \leq \infty$

$\|\underline{x}\|_q \leq \|\underline{x}\|_p$

Ex: $\max_i |x_i| \leq \sum_i |x_i|$

$\|\underline{x}\|_\infty \leq \|\underline{x}\|_1$

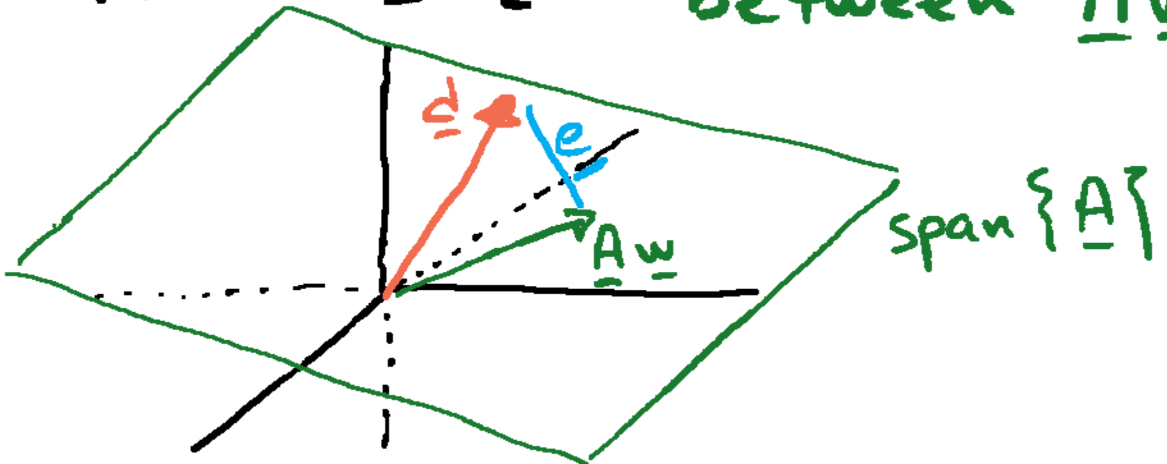
Different norms are used to encourage different attributes

- $\|\underline{x}\|_1$ sparse solutions
- $\|\underline{x}\|_\infty$ constant magnitude solutions
- $\|\underline{x}\|_2$ minimize squared error

Least-Squares Problem

$$\min_{\underline{w}} \|\underline{A}\,\underline{w} - \underline{d}\|_2^2$$

minimize squared error between $\underline{A}\underline{w}$ and $\underline{d}$



$\underline{e} = \underline{A}\underline{w} - \underline{d}$