$$\left[(x_1+y_1)^2 + (x_2+y_2)^2 + \cdots + (x_n+y_n)^2\right]^{1/2}$$

A norm is a thing that satisfies properties ① – ④.

The text defaults to the Euclidean norm.

$x \in \mathbb{R}^n \qquad x = (x_1, x_2, \ldots, x_n)$

$$\underbrace{\|x\| = \left[x_1^2 + x_2^2 + \cdots + x_n^2\right]^{1/2}}_{\text{"Euclidean norm"}}$$

"distance from zero vector"

1) $\|x\| \geq 0$

2) $\|x\| = 0 \iff x = 0$

3) $\|\alpha x\| = |\alpha| \, \|x\| \qquad \alpha \in \mathbb{R}$

4) $\|x+y\| \leq \|x\| + \|y\|$ "triangle inequality"

1) – 4) describe norms

There is a connection between norms and inner products

$$x^T x = x_1 \cdot x_1 + x_2 \cdot x_2 + \dots + x_n \cdot x_n$$

$$= x_1^2 + \dots + x_n^2$$

$$= \|x\|^2$$

$$\|x+y\|^2 = (x+y)^T(x+y)$$

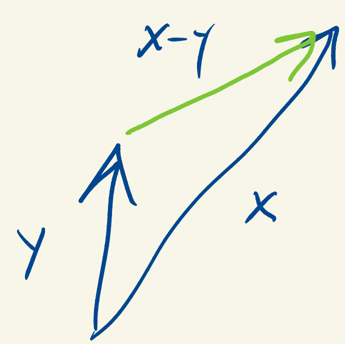
$$= x^Tx + x^Ty + y^Tx + y^Ty$$

$$= \|x\|^2 + 2x^Ty + \|y\|^2$$

"law of cosines in disguise"

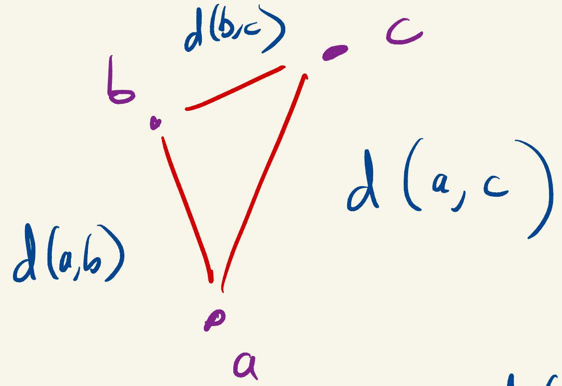
$$x^T \cdot y = \frac{1}{2}\left[\|x+y\|^2 - \|x\|^2 - \|y\|^2\right]$$

What should the distance from x to y be?



$$d(x,y) = \|x-y\|$$

"distance"



$$d(a,c) \overset{?}{\leq} d(a,b) + d(b,c)$$

$$\begin{aligned} d(a,c) &= \|a-c\| = \|a-b+b-c\| \\ &\leq \|a-b\| + \|b-c\| \\ &= d(a,b) + d(b,c) \end{aligned}$$

triangle ineq.

Feature vectors — all entries are 0 or 1, encoding whether some fact is false or true.

If x and y are binary feature vectors $0^2, 1, 1$

$0, -1, 1$

$||x-y||$

what is $d(x,y)$?

$$d(x,y) = (k)^{1/2}$$

k is the number of entries where x and y are different.

So $d(x,y)$ is a measure of dissimilarity between the vectors.

$$(\quad , \quad , \quad , \quad)$$

did grad from HS | college | post grad | age over 20

binary

$$x = (1, 1, 0, 17)$$

a 37 year old who graduated from HS and college.

$$\|x\| = \left(1^2 + 1^2 + 0^2 + 17^2\right)^{1/2}$$

$$= 17.05$$

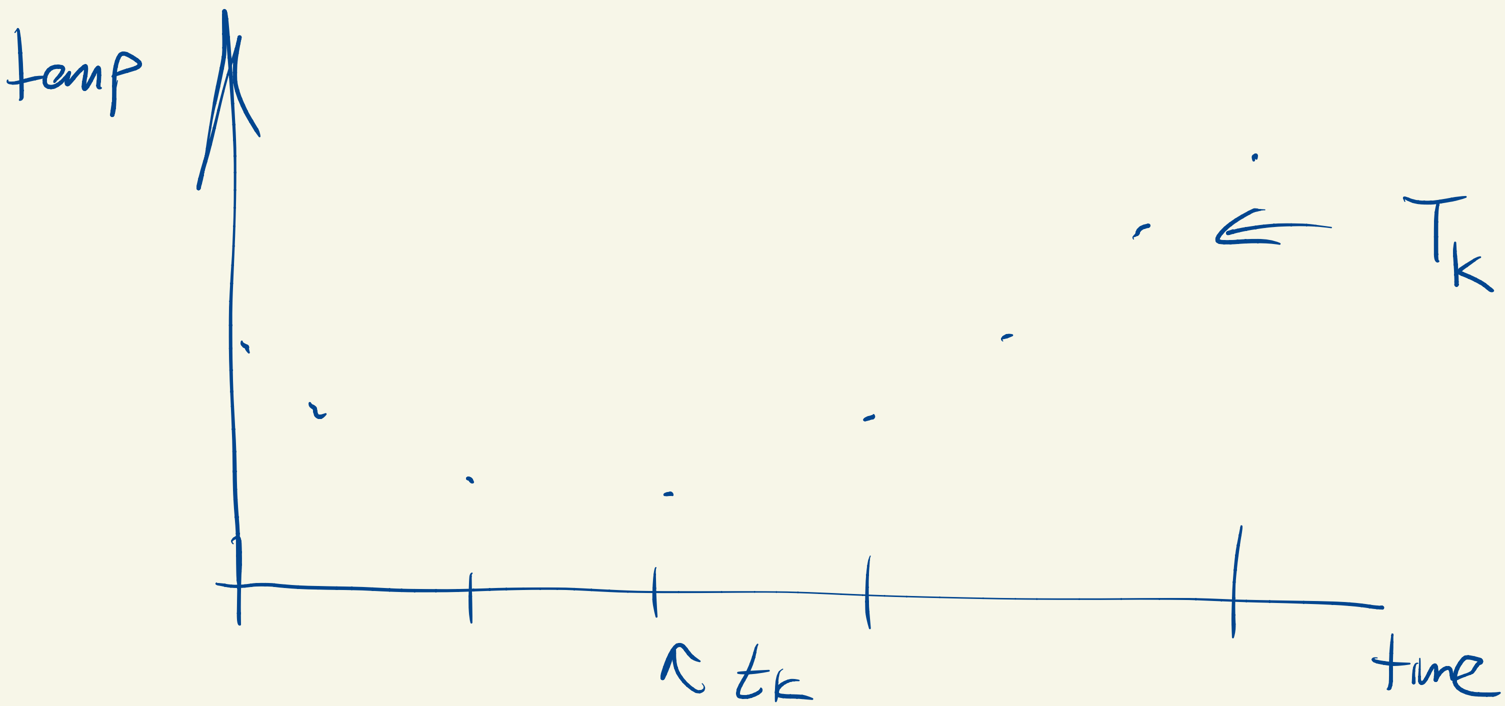
To avoid overly weighting the age we could use different units: decades over age 20.

Text problem 3.28 for an alternative approach using weights

$$x = (1, 1, 0, 1.7)$$

$$\|x\| = \left(1^2 + 1^2 + 0^2 + (1.7)^2\right)^{1/2} = 2.21$$

Time series



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

Temp T_k at time t_k $\quad k = 1, \ldots, n$

$$T = (T_1, T_2, \ldots, T_n) \quad \text{temp vector}$$

average temp $\quad \operatorname{avg}(T) = (T_1 + T_2 + \cdots + T_n) \frac{1}{n}$

$$= \frac{1}{n} \vec{1}^T T$$

$$\| \vec{1}_n \| = \Big[\underbrace{1^2 + 1^2 + \cdots + 1^2}_{n} \Big]^{1/2} = [\, n \,]^{1/2} = \sqrt{n}$$

The norm takes into account the size of the entries of the vector but also the number of entries

$$\mathrm{rms}(x) = \frac{\|x\|}{\sqrt{n}}$$

n

$m \cdots$

$$= \frac{(x_1^2 + x_2^2 + \cdots + x_n^2)^{1/2}}{n^{1/2}}$$

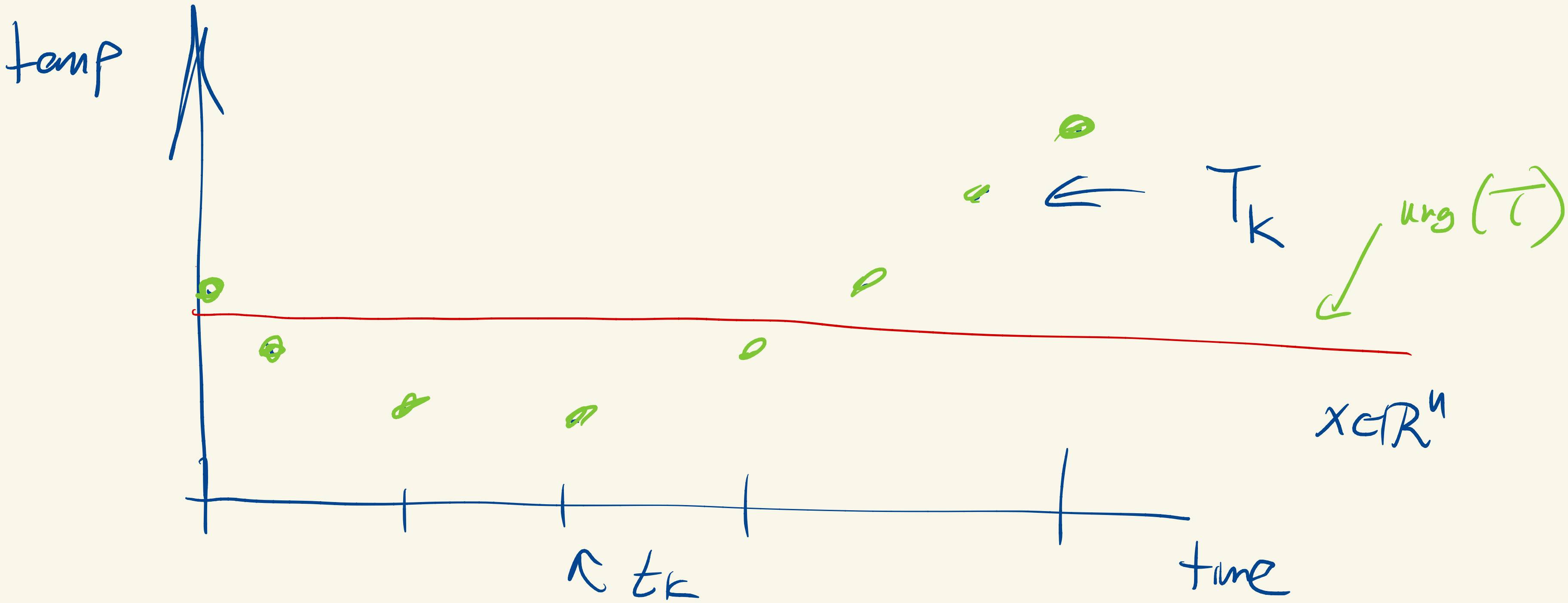
$$= \left[\frac{x_1^2 + x_2^2 + \cdots + x_n^2}{n}\right]^{1/2}$$

root mean square

$$\underbrace{\mathrm{rms}}_{\geq 0}(\vec{1}_n) = \frac{\|\vec{1}_n\|}{\sqrt{n}} = \frac{\sqrt{n}}{\sqrt{n}} = 1$$

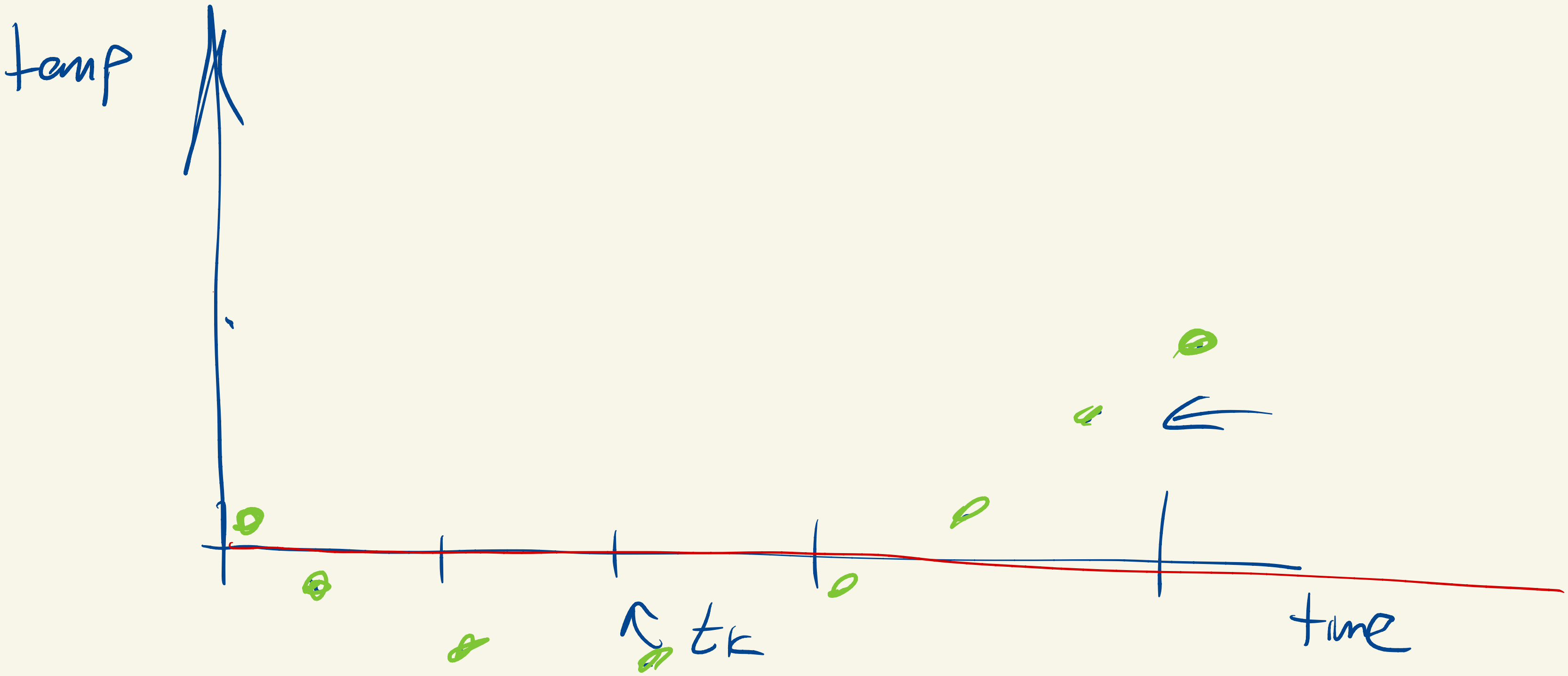
$x = (1, -1, -1, 1, -1, -1, -1, 1)$

$\mathrm{rms}(x) = 1$



rms (T) would be a kind of measure of the typical absolute value of an entry of T.

$$T - \mathrm{avg}(T)\vec{1}_n = (T_1 - \mathrm{avg}(T), T_2 - \mathrm{avg}(T), \ldots, T_n - \mathrm{avg}(T))$$

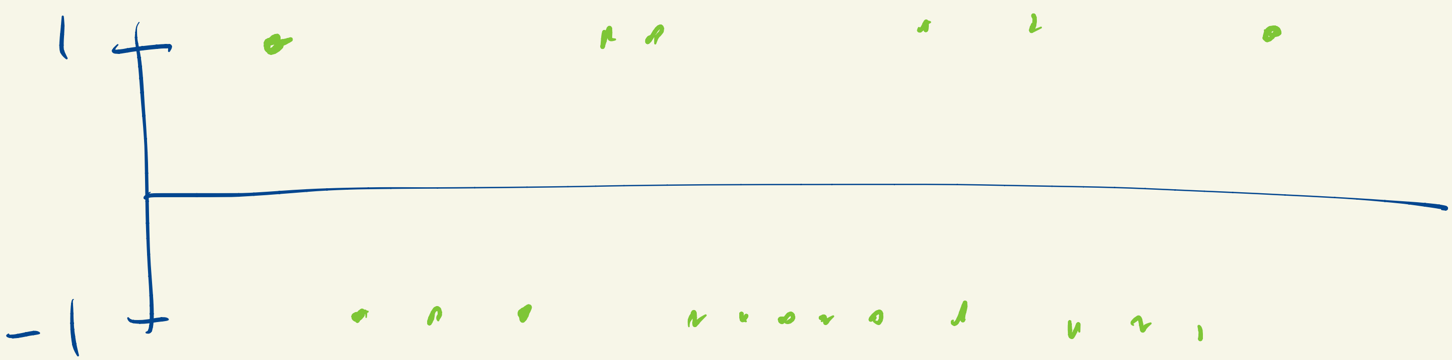


$$\frac{\|T - \mathrm{avg}(T)\vec{1}_n\|}{\sqrt{n}} = \mathrm{rms}(T - \mathrm{avg}(T)\vec{1}_n)$$

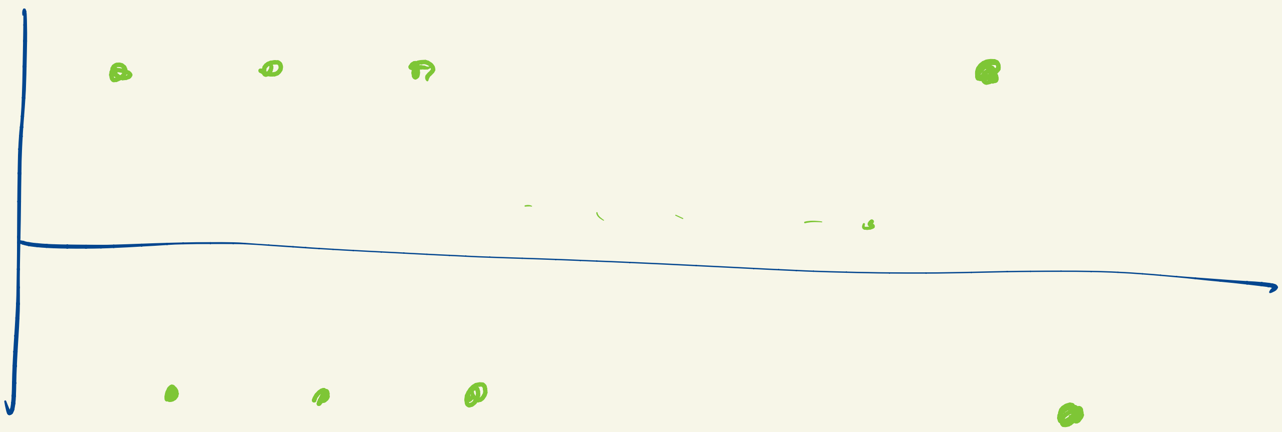
$$:= \mathrm{std}(T)$$

$$\mathrm{std}(\vec{1}_n) = 0$$

$$\mathrm{std}(x) = \left[\frac{\sum_{k=1}^{n} (x_k - \mathrm{avg}(x))^2}{n}\right]^{1/2}$$



std(



x

$$\mathrm{std}(x) = ? \qquad \mathrm{avg}(x) = 0$$

$$\begin{aligned}\mathrm{std}(x) &= \mathrm{rms}\left(x - \overbrace{\mathrm{avg}(x)}^{0}\,\mathbf{1}\right)\\ &= \mathrm{rms}(x)\\ &= \left[\frac{1^2 + (-1)^2 + \cdots + 1^2 + (-1)^2}{n}\right]^{1/2}\end{aligned}$$

$$= \left[\frac{n}{n}\right]^{1/2} = \sqrt{1} = 1$$