CPSC 536F Sept 12, 2025

- Remark [TL]: Is it really more efficient:

$$(1+\vec{x}\cdot\vec{x}')^2 \quad \text{vs.}$$

$$\Phi(\vec{x})\cdot\Phi(\vec{x}'), \quad \text{where}$$

$$\Phi(\vec{x})=(1,\sqrt{2}x_1,\sqrt{2}x_2,x_1^2,\sqrt{2}x_1x_2,x_2^2)$$

?

- Lemma: Let $\varphi^{test}=\varphi\in\mathbb{R}^6$ $C, G\subset\mathbb{R}^6$. Then to know if

$$\Big\| \varphi - \overbrace{\frac{1}{|C|} \sum_{c \in C} c}^{\text{average/centre of mass cows}} \Big\|^2_{L^2}$$

$$\leq \Big\| \varphi - \underbrace{\frac{1}{|G|} \sum_{g \in G} g}_{\text{average ghost goats}} \Big\|^2_{L^2}$$

it suffices to know $\vec{a} \cdot \vec{b}$ where $\vec{a}, \vec{b}$ range over all of $\{\varphi\} \cup C \cup G \subset \mathbb{R}^6$

Today: More examples of kernel functions / positive (semi)def

Remark [TL]:

- If $\left(1 + \vec{x} \cdot \vec{x}'\right)^d$

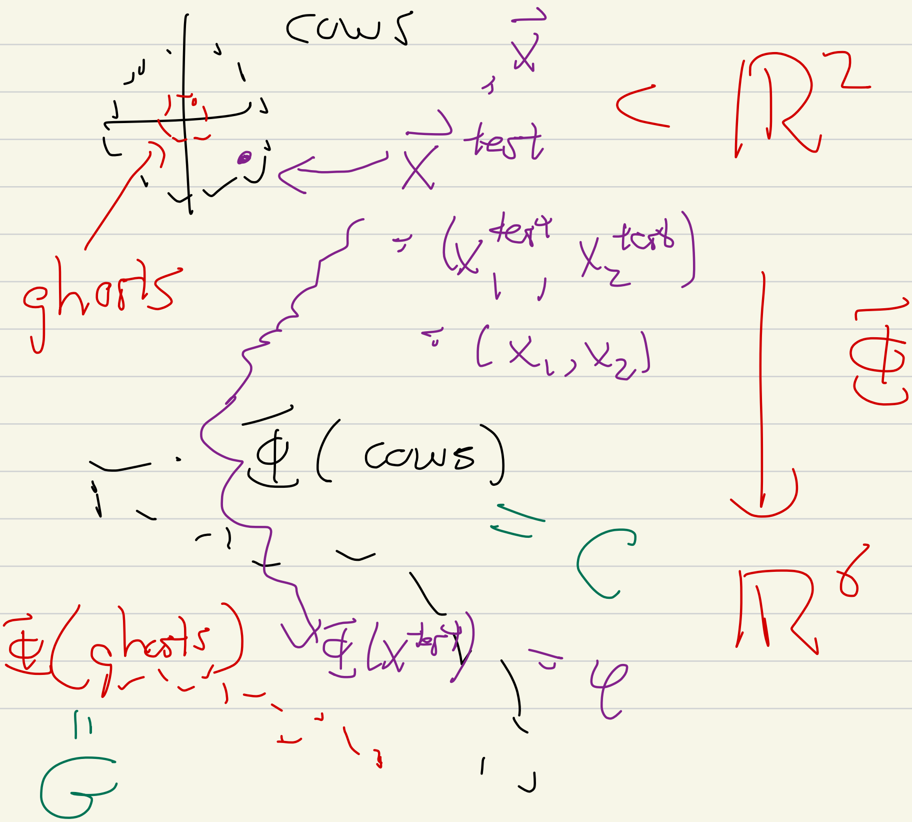
  $\vec{x}, \vec{x}' \in \mathbb{R}^n$, n larger $(n \gg 2)$

  $d \in \mathbb{N} = \{1, 2, \ldots\}$ $(d \gg 2)$

- There are conceptual reasons to with kernel functions: symmetric, positive (semi) definite

- Hopefully: this will give explanations as to why an algorithm may work poorly, degenerate, ...

Maybe we can see why regression can "break" but ridge regression can solve things.



Pf: First Exercise on homework

(with $\mathbb{R}^1 \longrightarrow \mathbb{R}^m$, any m.)

$$\left\{ \vec{a} \cdot \vec{b} \;\middle|\; \vec{a}, \vec{b} \in \{\varphi\} \cup C \cup G \right\}$$

So if

$$C = \Phi(\text{cows})$$

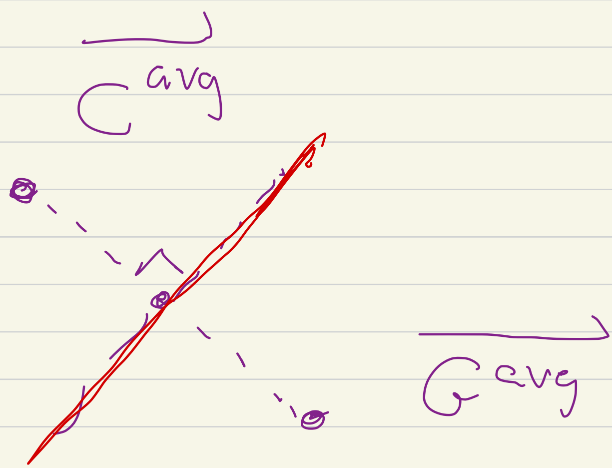
$$G = \Phi(\text{ghosts})$$

$$\varphi = \Phi(\vec{x}_{\text{test}})$$

$$\vec{a} \cdot \vec{b} = \Phi(\vec{u}) \cdot \Phi(\vec{v}) = \left\{ \begin{array}{c} \text{simpler} \\ \text{formula} \end{array} \right\}$$

Also

$$\overrightarrow{C^{avg}} = \frac{1}{|C|} \sum_{c \in C} c$$



<u>Thm</u>: If $K \in \mathbb{R}^{n \times n}$ is a real $n \times n$ matrix s.t. $K^T = K$ (i.e. $K_{ij} = K_{ji}$, i.e. K is symmetric), then K has an orthonormal eigenbasis, i.e. there are $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ s.t.

(1) $\vec{v}_i \cdot \vec{v}_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$

(2) $K\vec{v}_i = \lambda_i \vec{v}_i$ where $\lambda_i \in \mathbb{R}$

Example:

$$K = \begin{bmatrix} d_1 & & & 0 \\ & d_2 & & \\ & & \ddots & \\ 0 & & & d_n \end{bmatrix}$$

$$K = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix} \in \mathbb{R}^{2\times 2}$$

$$\begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix} \underbrace{\begin{bmatrix} 1 \\ 0 \end{bmatrix}}_{\overline{e}_1} = d_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} d_2$$

$\vec{e}_2$

$\vec{e}_i$ = standard basis vector.

$$\begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}^3 = \begin{bmatrix} d_1^3 & 0 \\ 0 & d_2^3 \end{bmatrix}$$

$$p(x) = c_0 x^n + c_1 x^{n-1} + \dots + c_n$$

$c_0, \dots, c_n \in \mathbb{R}$,

$$p\left(\begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}\right)$$

$$= \begin{bmatrix} p(d_1) & 0 \\ 0 & p(d_2) \end{bmatrix}$$

$t \in \mathbb{R}$

$$e^{t\begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}} =$$

$$\left[e^A \overset{def}{=} I + A + \frac{A^2}{2} + \frac{A^3}{6} + \frac{A^4}{4!} + \cdots \right]$$

$$p\left(\begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}\right)$$

$$= c_0 \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}^n + c_1 \begin{bmatrix} \cdot \end{bmatrix}^{n-1}$$

$$+ \cdots + c_{n-1} \begin{bmatrix} \cdot \end{bmatrix} + \underbrace{c_n \begin{bmatrix} \cdot \end{bmatrix}^0}_{c_n I}$$

$$p(A) \stackrel{\text{def}}{=} c_0 A^n + c_1 A^{n-1} + \cdots + c_{n-1} A + \underbrace{}_{c_n I}$$

Now:

$$\begin{bmatrix} a & b \\ b & a \end{bmatrix},$$

$$\begin{bmatrix} a & b \\ b & a \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} a & b \\ b & a \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \overbrace{(a+b)}^{\lambda_1}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

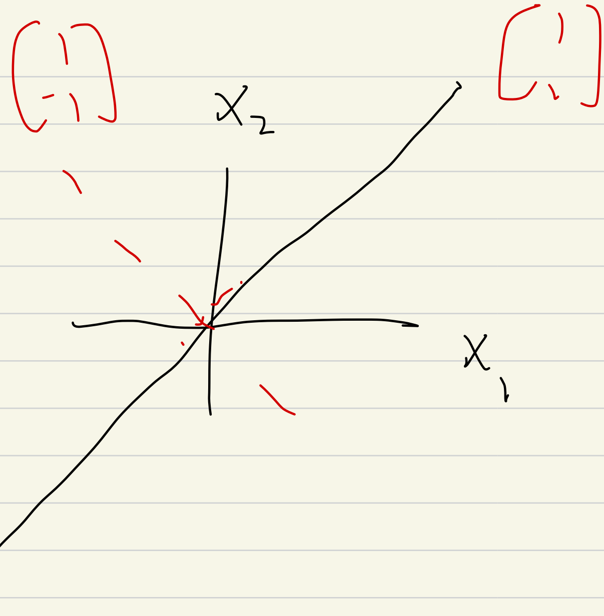
$$\begin{bmatrix} a & b \\ b & a \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \underbrace{(a-b)}_{\lambda_2}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\mathrm{Tr}\begin{pmatrix} a & b \\ b & a \end{pmatrix} = 2a$$

$$= \lambda_1 + \lambda_2$$

$a=0,\ b=1$

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$



$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ x_1 \end{bmatrix}$$

$$\underbrace{\begin{bmatrix} a & b \\ b & a \end{bmatrix}}_{K}\underbrace{\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}}_{M} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} a+b & \\ & a-b \end{bmatrix}$$

$$K \cdot M = M\Lambda,$$

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

$$K = M \Lambda M^{-1}$$

$$K^2 = M \Lambda M^{-1} M \Lambda M^{-1}$$

$$= M \Lambda^2 M^{-1}$$

$$p(K) = M\, p(\Lambda) M^{-1}$$

$$A_1 = M B_1 M^{-1}$$

$$A_2 = M B_2 M^{-1}$$

$$A_3 = M B_3 M^{-1}$$

$$A_4 = M B_4 M^{-1}$$

$$A_1 A_2 A_3 A_4$$

$$= M B_1 B_2 B_3 B_4 M^{-1}$$

$$K\begin{bmatrix} | & | & \\ \vec{v}_1 & \vec{v}_2 & \cdots \\ | & | & \end{bmatrix} = \begin{bmatrix} | & | & \\ \vec{v}_1 & \vec{v}_2 & \cdots \\ | & | & \end{bmatrix} \overbrace{\Lambda}^{\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \ddots \end{bmatrix}}$$

If $\vec{v}_1, \ldots, \vec{v}_n$ are orthonormal,

then

$$K = \sum_{i=1}^{n} \lambda_i \vec{v}_i \vec{v}_i^{T}$$

Check: $K\vec{v}_j = \left(\sum_{i=1}^{n} \lambda_i \vec{v}_i \vec{v}_i^{T}\right)\vec{v}_j$

$= \lambda_j \vec{v}_j$

If $K^T = K$

$$K U = U \underbrace{\Lambda}_{\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}}$$

Fact K is positive semidefinite

$$Q(\vec{v}) = \vec{v} \cdot (K \vec{v})$$
$$= (K\vec{v}) \cdot \vec{v} \geq 0$$

for all $\vec{v}$ IFF $\lambda_i \geq 0$ for all i

$$K = \begin{bmatrix} 2 & & & & \\ & 3 & & & \\ & & 1 & & \\ & & & 0 & \\ & & & & 4 \end{bmatrix}$$

Then:

$$e^{Kt} = U \underbrace{\begin{bmatrix} e^{2t} & & & & \\ & e^{3t} & & & \\ & & e^{t} & & \\ & & & 1 & \\ & & & & e^{4t} \end{bmatrix}}_{\text{all } \lambda\text{'s} > 0} \vec{U}^{-1}$$

iff positive definite matrix

$$\frac{d}{dt} \vec{y}(t) = K \vec{y}(t)$$