Math Foundations for ML

10-606

Geoff Gordon

Tips for written homework

- For HW1, you'll need to submit the written part through Gradescope
- If you're registered, please access Gradescope via Canvas: open Canvas, select Gradescope from the left navigation column
  - this lets us link your Gradescope account to Canvas so you get credit
  - if you're not registered, should still be able to access Gradescope directly
- Two options for preparing submission
  - handwrite and scan
  - type using a markup language

Handwrite and scan

- Handwrite legibly!
- Use a scanning app on your phone (see suggestions on course website)
  - don’t just take a photo; this will result in skewed, poor-contrast, badly cropped images that are hard for the TAs to work with
- Upload PDF to Gradescope

Markup languages

- Two most common are LaTeX and Markdown; both work well but need setup
  - Both can produce PDFs to submit on Gradescope
- For lecture notes, I use the VSCode editor (Markdown is built in) with Markdown+Math (enables LaTeX math) and Markdown Extended (better CSS)

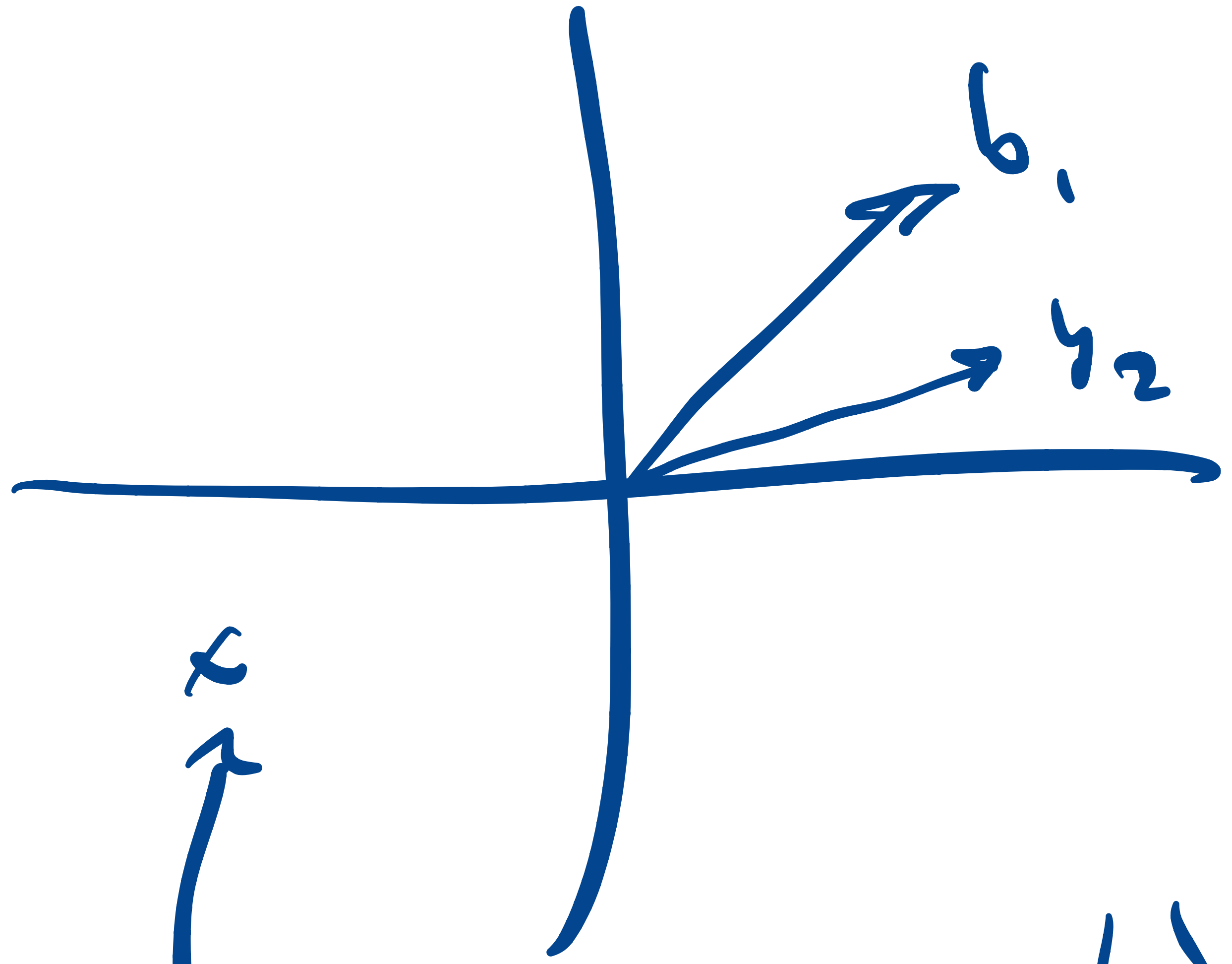
```
## Complete spaces

Above we described how to think of matrices
or functions as vectors in a vector space. We
also described how to upgrade a vector space
to an inner product space by defining an inner
product $\langle x, y\rangle$. For example,
- A useful inner product for matrices is $$
\langle X, Y\rangle = \sum_{i=1,\,j=1}^{i=m,
\,j=n} X_{ij}Y_{ij} = \mathrm{tr}(X^TY) =
\mathrm{tr}(YX^T)$$
```

Complete spaces

Above we described how to think of matrices or functions as vectors in a vector space. We also described how to upgrade a vector space to an inner product space by defining an inner product $\langle x, y\rangle$. For example,

- A useful inner product for matrices is

$$\langle X, Y\rangle = \sum_{i=1,\,j=1}^{i=m,\,j=n} X_{ij}Y_{ij} = \mathrm{tr}(X^TY) = \mathrm{tr}(YX^T)$$



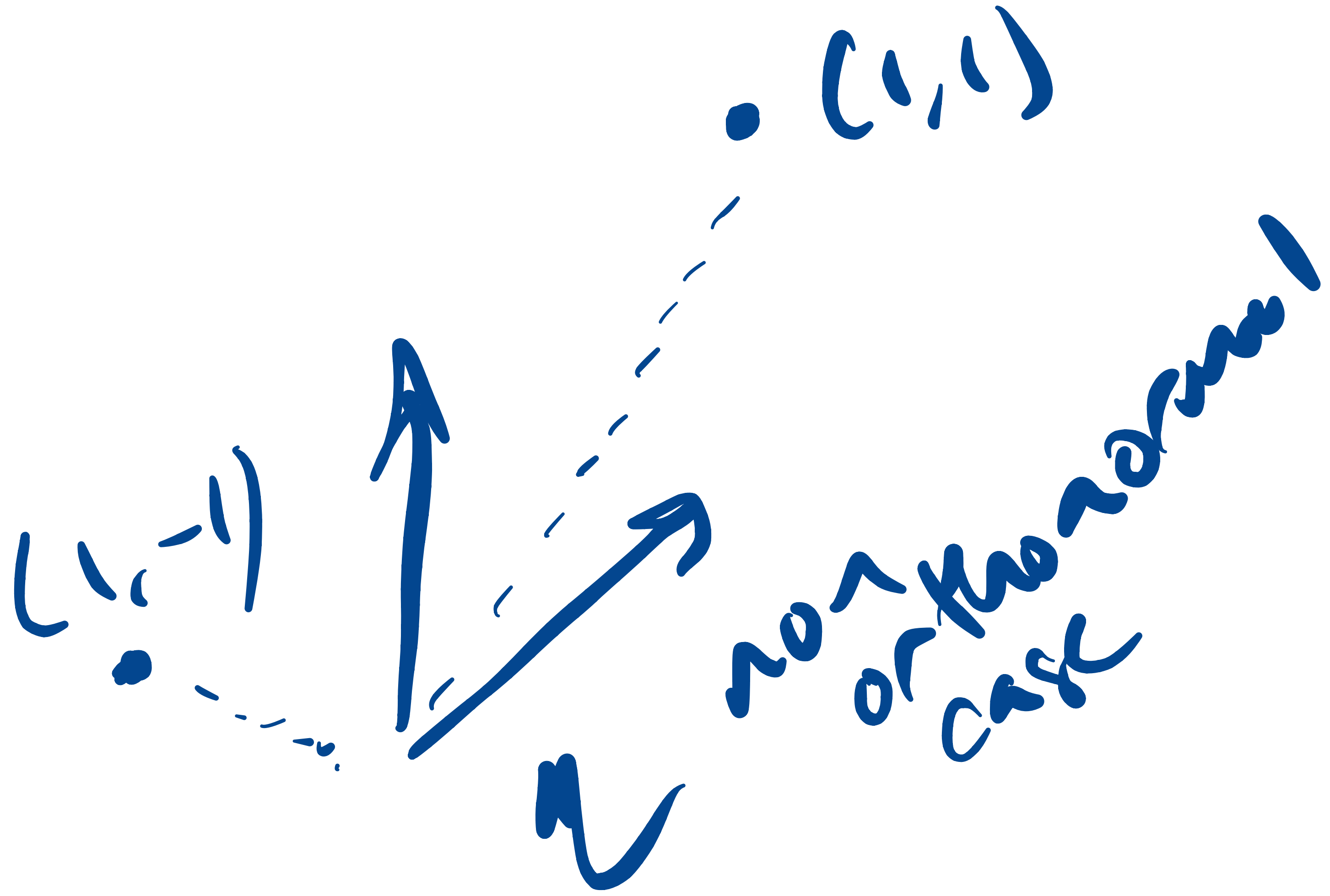
x

$b_1 - b_2 \rightarrow \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

$\{ w_1 b_1 + w_2 b_2 \mid w_1, w_2 \in \mathbb{R} \}$

↳ span

$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \cancel{\begin{pmatrix} 0 \\ 1 \end{pmatrix}} \quad \begin{pmatrix} 1 \\ 1 \end{pmatrix}$



orthonormal

↳ $\langle x, y \rangle = u \cdot v$

$x \downarrow u \quad y \downarrow v$

Exercise: equivalent vector spaces

- Are these vector spaces the same? $R^{m \times 1}$ $R^{1 \times m}$ R^m
  - A: yes, they're the same
  - B: no, they're different
  - C: they're different but equivalent

Exercise: basis

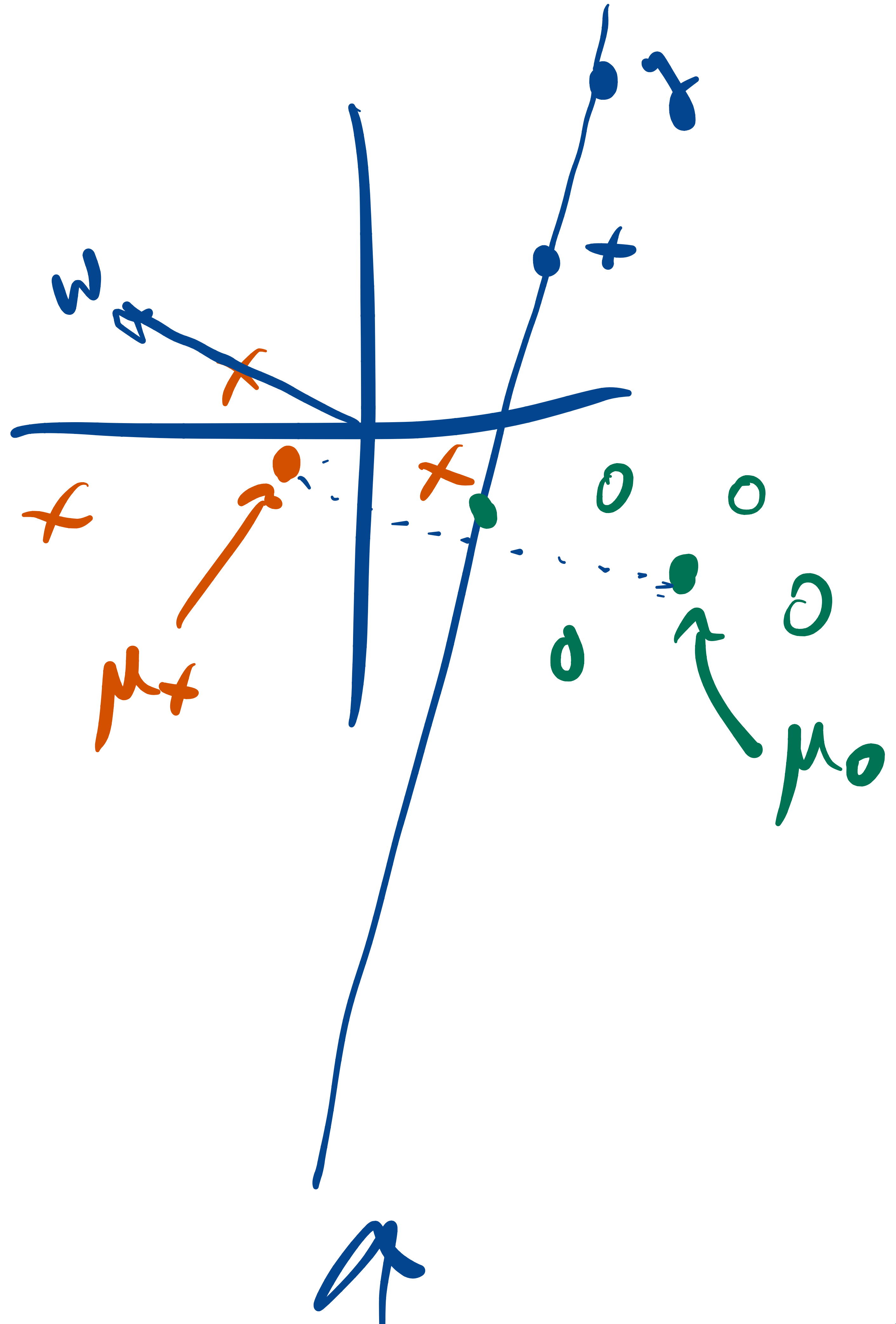
Go to text box on Canvas to answer

- Consider the vector space of degree-2 polynomials in a real variable x with the basis 1, x, $2x^2 - 1$ (first three Chebyshev polynomials)

- What is the representation of x^2 in this basis?

$$\frac{1}{2}\cdot 1 + 0 + \frac{1}{2}(2x^2-1) \rightarrow \begin{pmatrix} 1/2 \\ 0 \\ 1/2 \end{pmatrix}$$

- What is the representation of $(x-1)^2$?



$w \in \mathbb{R}^2 \qquad b \in \mathbb{R}$

x if $w \cdot x > b$

o if $w \cdot x < b$

$f(x) = w \cdot x - b$

$\{ x \mid f(x) = 0 \}$

decision surface

$w \cdot x - \cancel{b} - (w \cdot y - \cancel{b}) = 0$

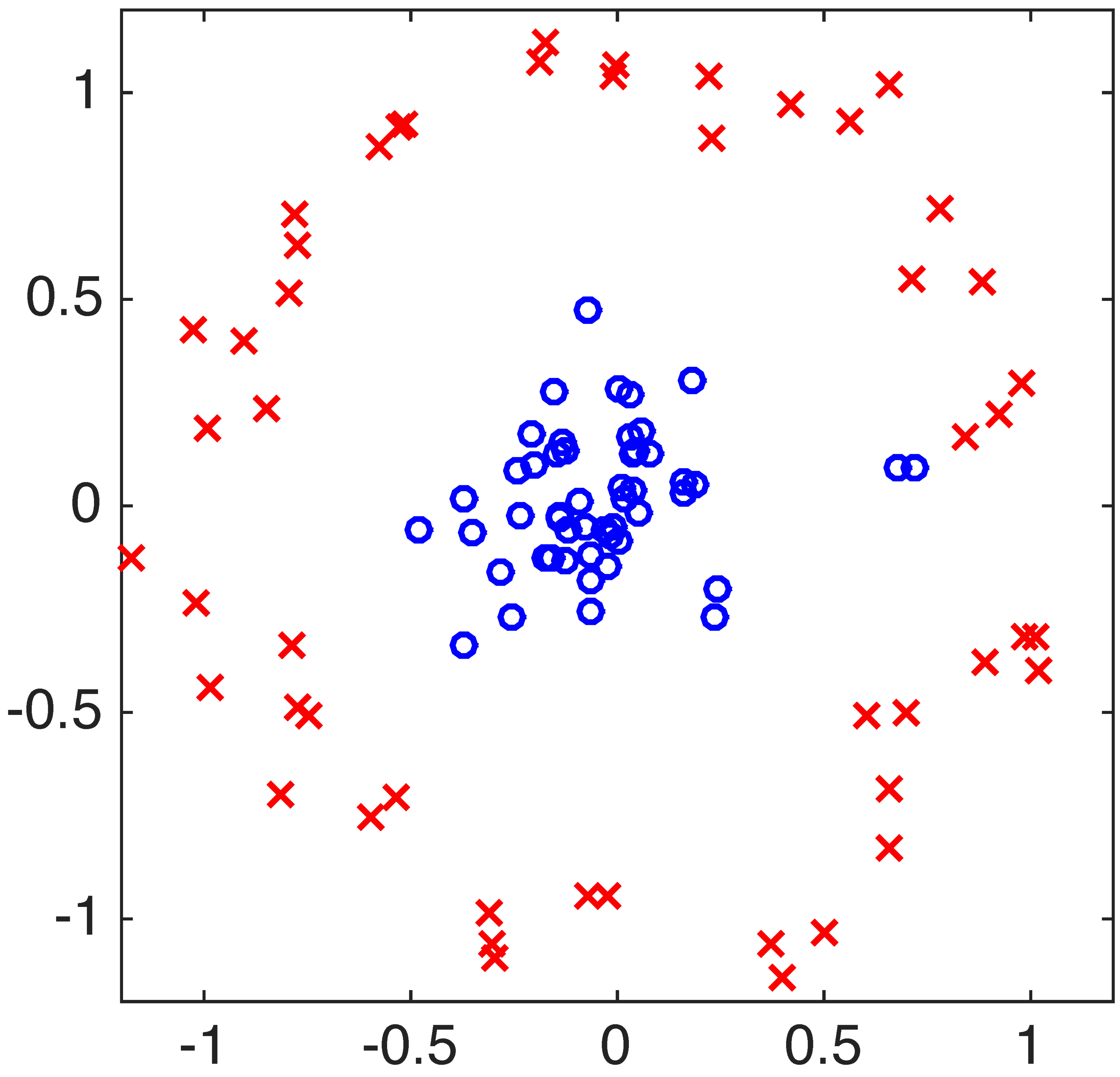
$w \cdot (x - y) = 0$

$kw \cdot w - b = 0$

$k = \frac{b}{\|w\|^2}$

$kw = \frac{w}{\|w\|} \boxed{\frac{b}{\|w\|}}$ dist from origin

Feature transforms

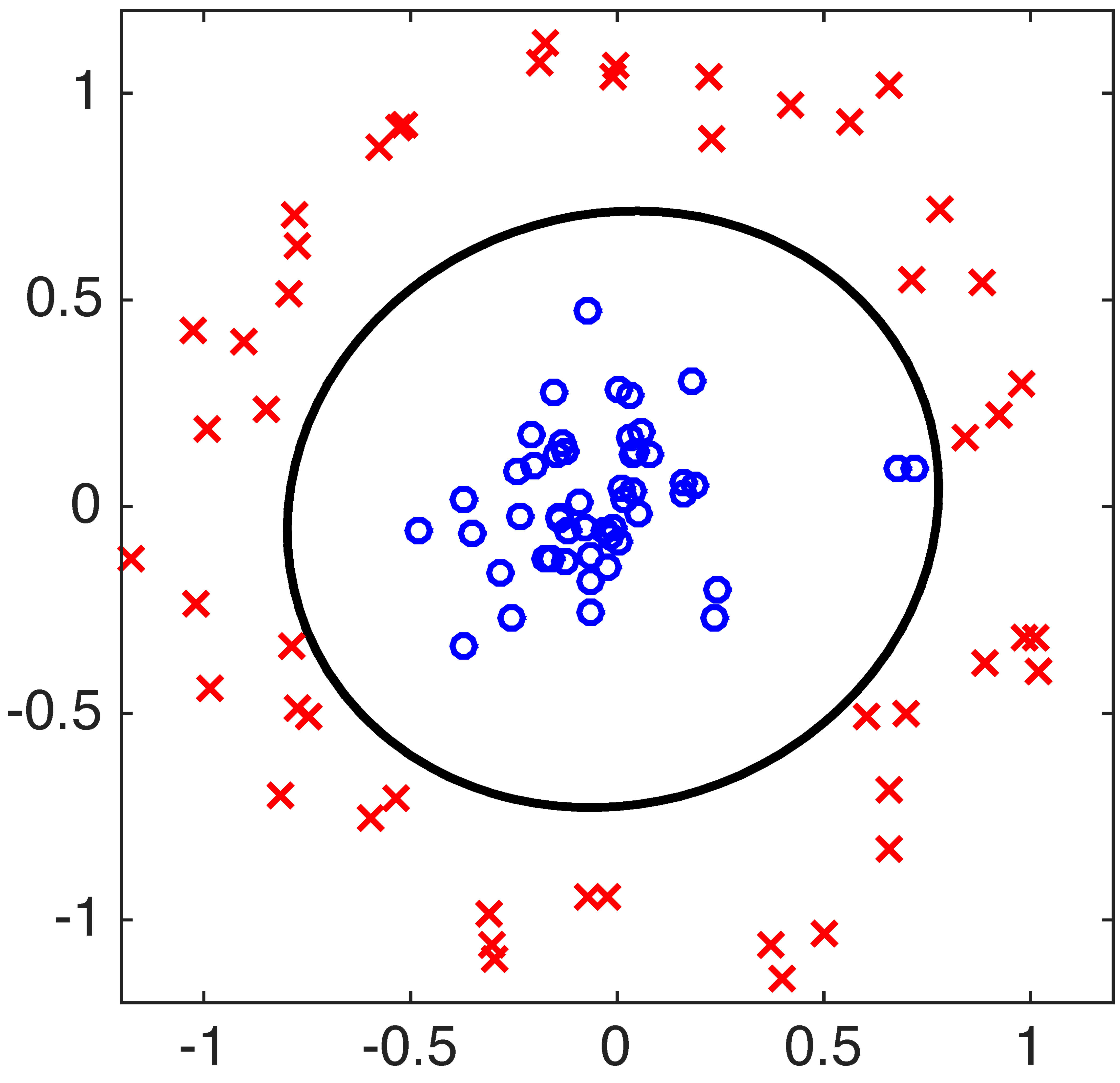


$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \xrightarrow[\phi]{} \begin{bmatrix} x_1 \\ x_2 \\ x_1 x_2 \\ x_1^2 \\ x_2^2 \end{bmatrix}$$

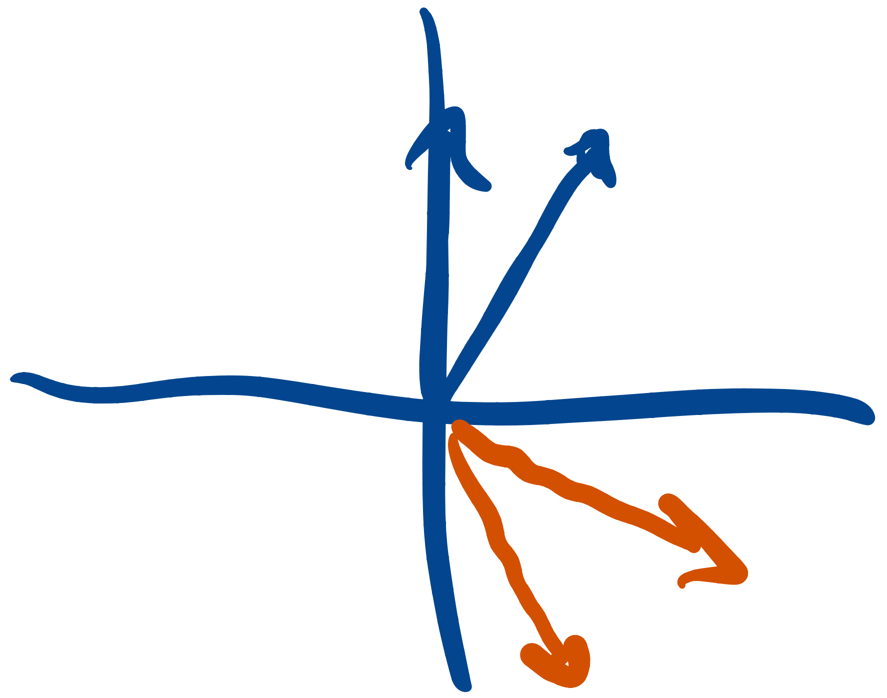
$$w \cdot \phi(x) - b$$

$$w_1 x_1 + w_2 x_2 + w_3 x_1 x_2 + w_4 x_1^2 + w_5 x_2^2 - b = 0$$

Feature transforms



$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \to \begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix}$$



$L \in U \to V$

basis   basis

$b_1 \dots b_n$ $c_1 \dots c_m$

$$\begin{tikzcd} U \arrow[r, "L"] \arrow[d] & V \arrow[d] \\ \mathbb{R}^n \arrow[r, "l"] & \mathbb{R}^m \end{tikzcd}$$

$\cancel{L(x)}$

$L(b_j) \in V$

$$L(b_j) = l_{1j} c_1 + l_{2j} c_2 + \dots l_{mj} c_m$$

$$\begin{pmatrix} l_{11} & \dots & l_{n1} \\ \vdots & \ddots & \vdots \\ l_{1m} & \dots & l_{nm} \end{pmatrix}$$

$\hookrightarrow \in \mathbb{R}^n \to \mathbb{R}^m$

$$f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x+y \\ 2x+2y \end{pmatrix} \to \text{range} = \text{span}\begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$$\to \text{null} = \text{span}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

rank + nullity = n

Adjoint

f^*

$$\left\langle \begin{pmatrix} u \\ v \end{pmatrix}, f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) \right\rangle$$

$$= \left\langle f^*\left(\begin{pmatrix} u \\ v \end{pmatrix}\right), \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle$$

$$\begin{pmatrix} -2 \\ 1 \end{pmatrix} \in \text{null}(f^*) \qquad f^*\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u+2v \\ u+2v \end{pmatrix}$$

Inverse

$$\begin{pmatrix} u \\ v \end{pmatrix} = f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right)$$

$$f^{-1}\left(\begin{pmatrix} u \\ v \end{pmatrix}\right) = \begin{pmatrix} x \\ y \end{pmatrix}$$

$\text{range}(f) \perp \text{null}(f^*)$

Exercise: Gaussian elimination

- Suppose
  - $x + y + z = 3$
  - $2x + y = 5$
  - $-x + y - 2z = 4$
- What are x, y, z?

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 0 \\ -1 & 1 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3 \\ 5 \\ 4 \end{pmatrix}$$

$$x + y + z = 3$$

$$2x + y = 5$$

$$-x + y - 2z = 4$$

$$0 + 2y - z = 7$$

$$0 + -y - 2z = -1$$

$$0 + 0 - 5z = 5$$