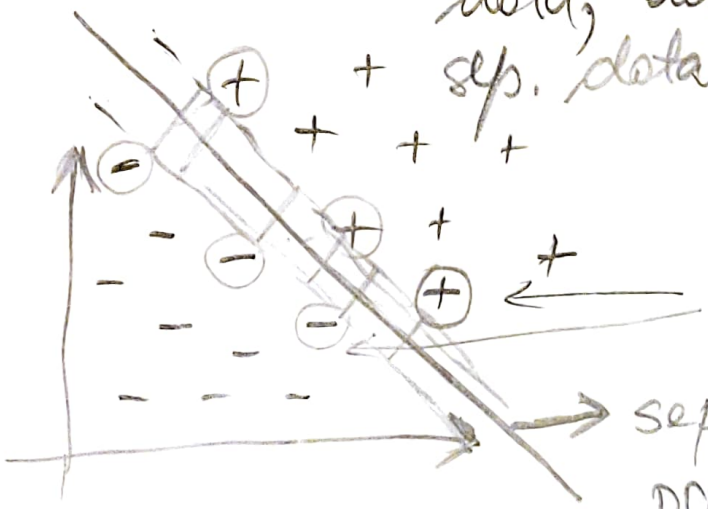
SVM → work only with 2 classes
→ they produce a line (or hyperplane)
⇓
work best for linearly separable data; but can work for non-lin. sep. data by allowing some errors



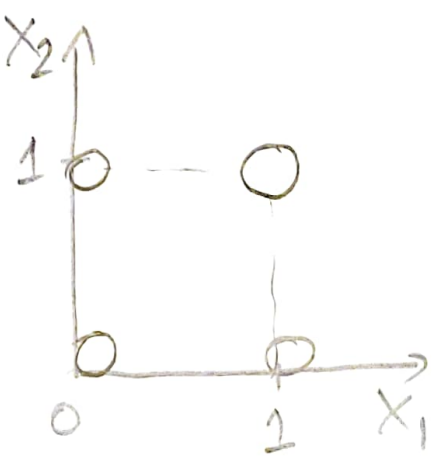
(!) SVM vs. perceptron
→ gives a line that separate + ≠ −
→ gives best line that separates + ≠ −
→ with largest margin

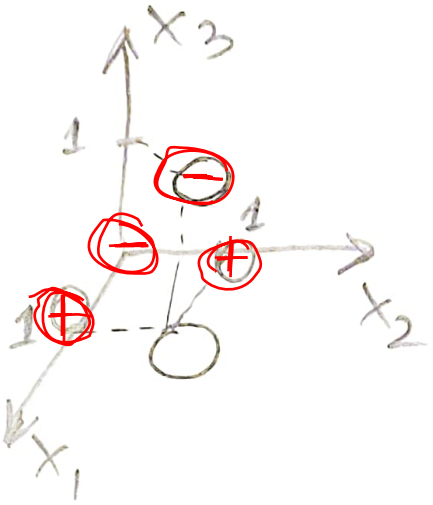
(!) SVMs are used in conjunction with kernels (≡ mapping fct.):
Step I Apply kernel to data to map it in a space where it is lin. separable
Step II Apply SVM to find the (optimum) separation hyperplane

Kernel ex.: polynomial; goussian; ...

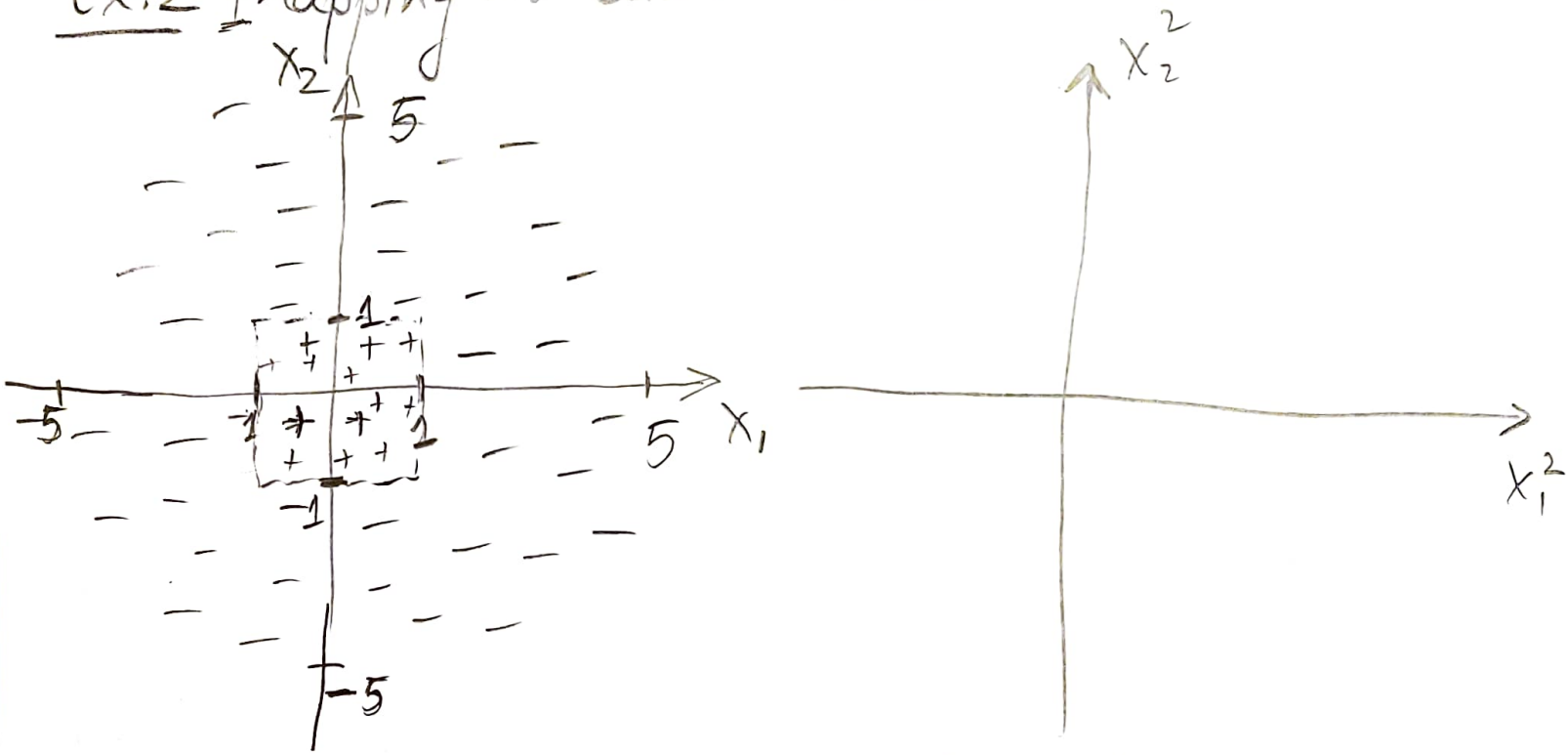
Ex.1 Mapping to higher dim.

x_1	x_2	xor	$x_1 * x_2$ ($= x_3$)
0	0	0 ⊖	0
0	1	1 ⊕	0
1	0	1 ⊕	0
1	1	0 ⊖	1





Ex.2 Mapping to same dimension



$$(x_1, x_2) \longrightarrow (x_1^2, x_2^2)$$

⊕ $(\frac{1}{2}, \frac{1}{2}) \longrightarrow$

⊖ $(3, 4) \longrightarrow$

⊕ $(-\frac{1}{2}, -\frac{1}{3}) \longrightarrow$

⊖ $(2, -3) \longrightarrow$