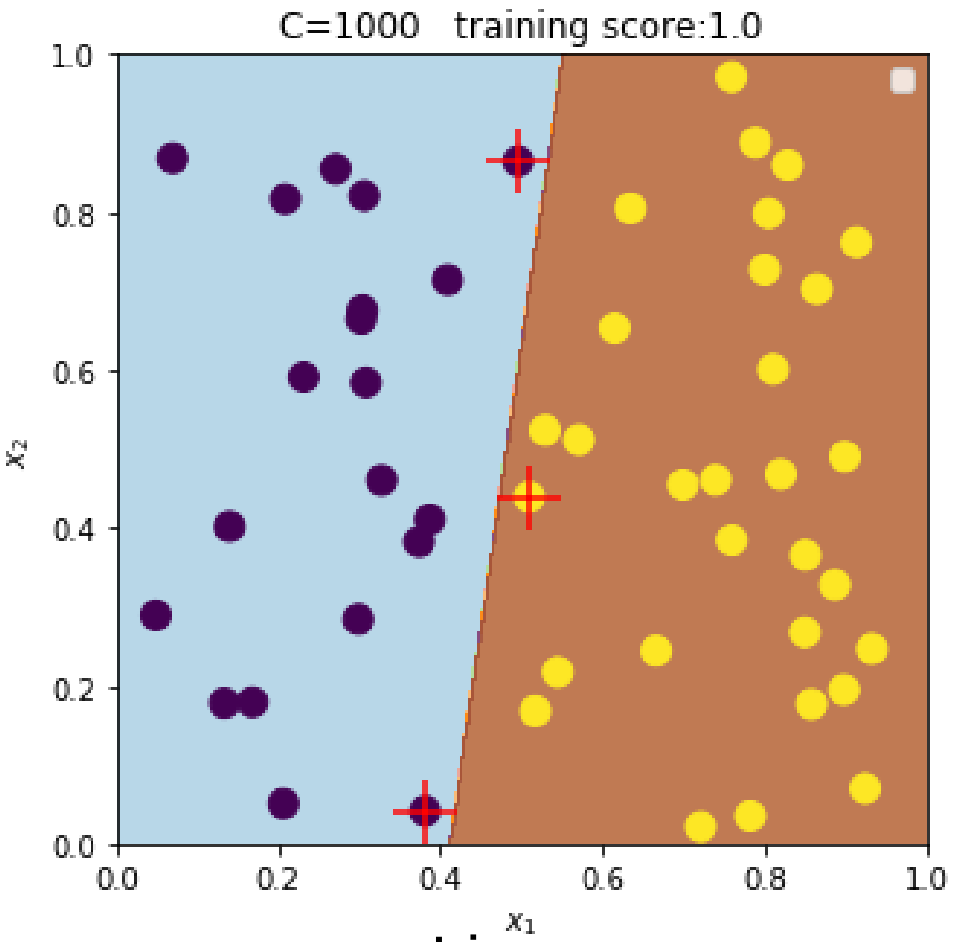
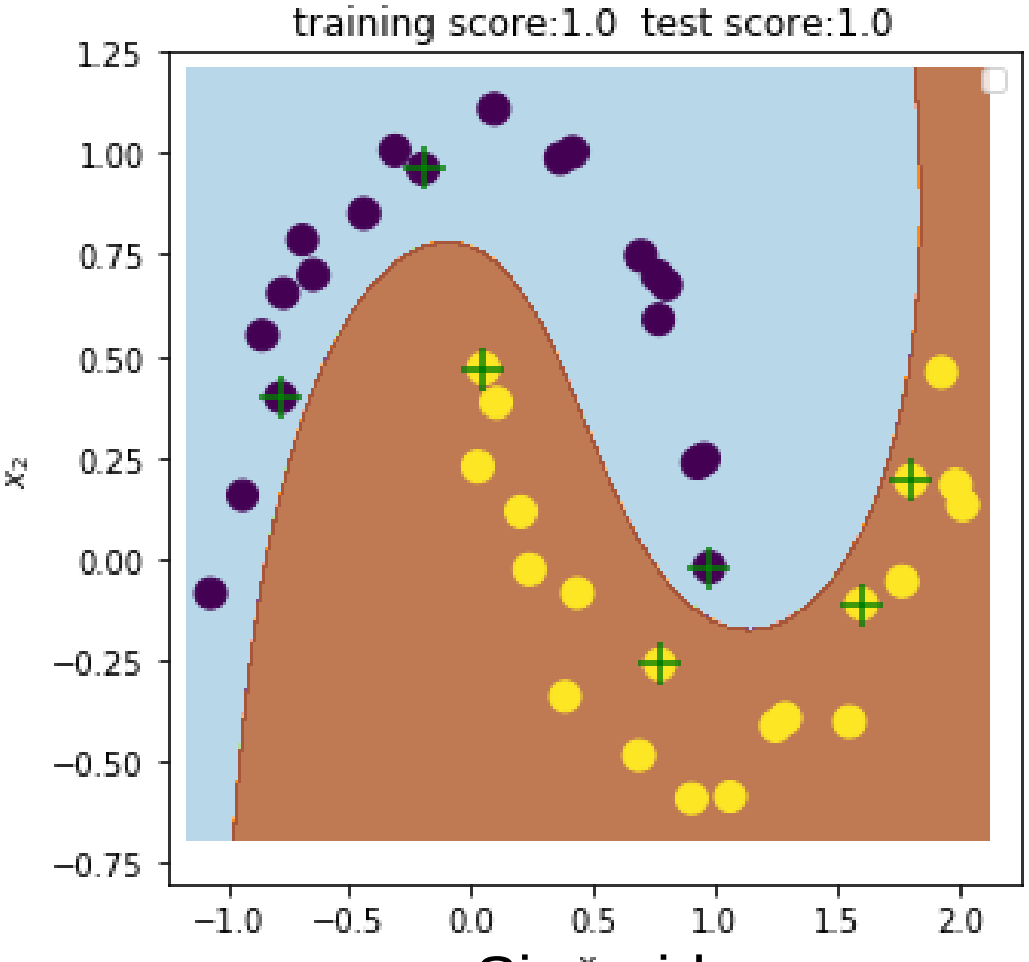
Lecture 6 – Part II
Support Vector Machines (SVM)

Support Vector Machines
Separable case



Linear



Sigmoid

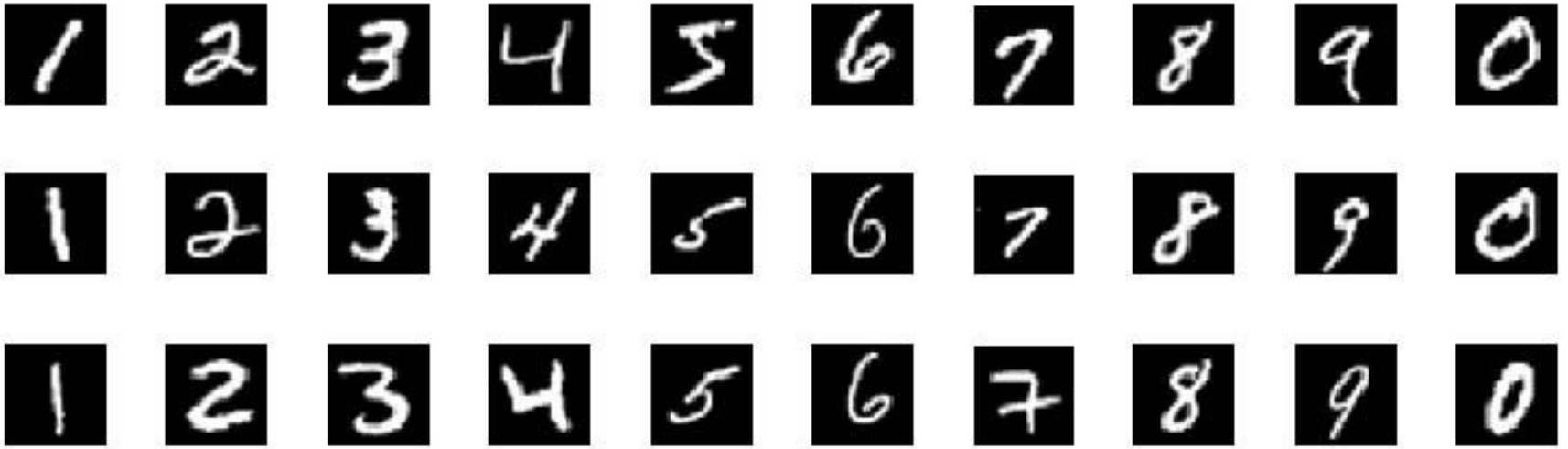
Separable case: SVM finds *the **best*** (linear) separation

SVM – separable case

SVM – separable case

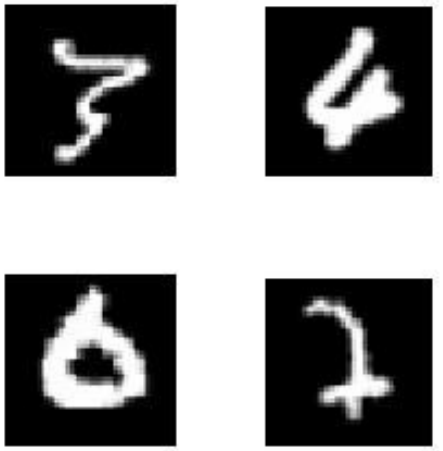
Support vectors

- The whole data:



- The Support Vectors:

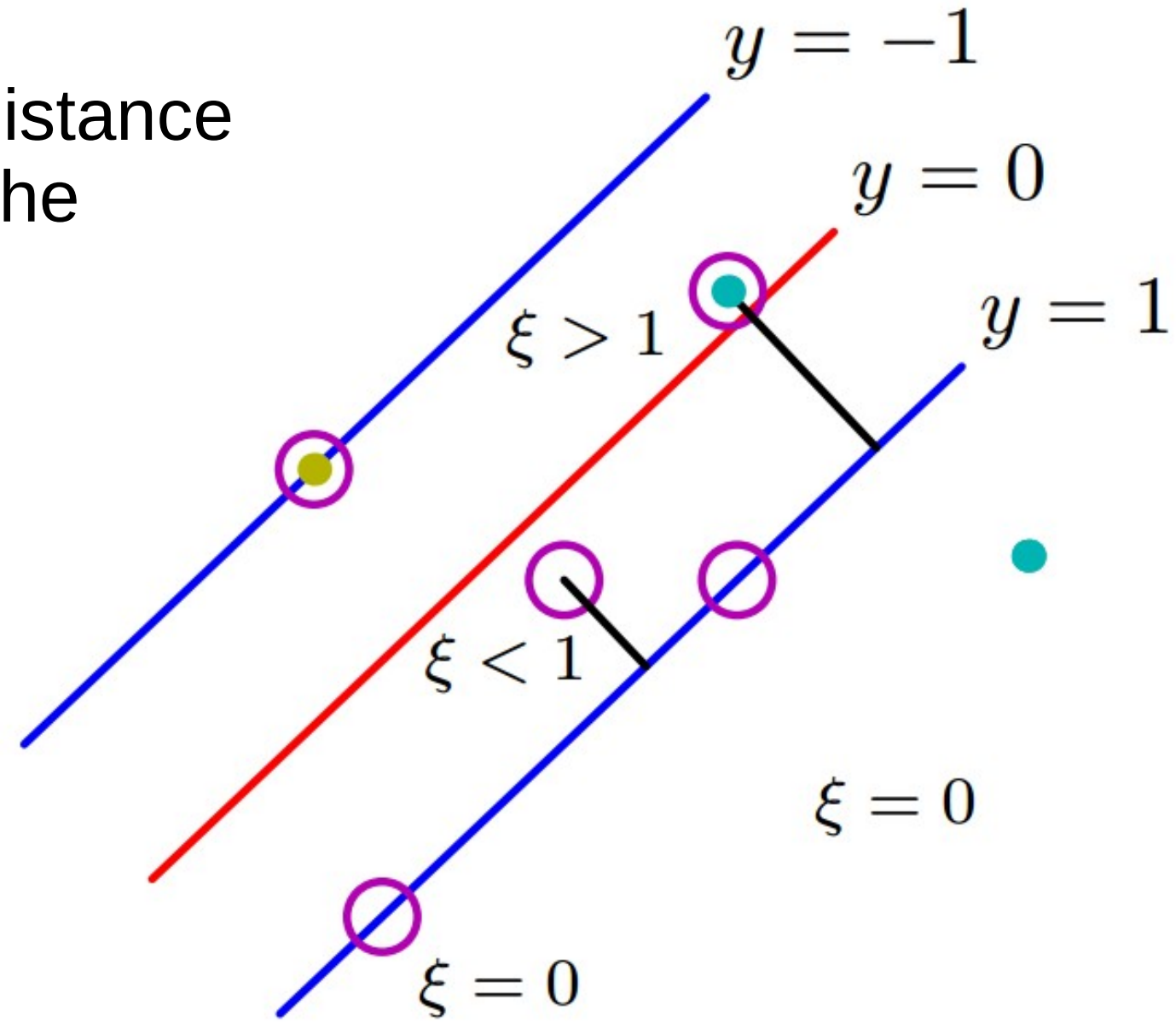
  (≃ the weirdos (outliers))



SVM – **non** separable case

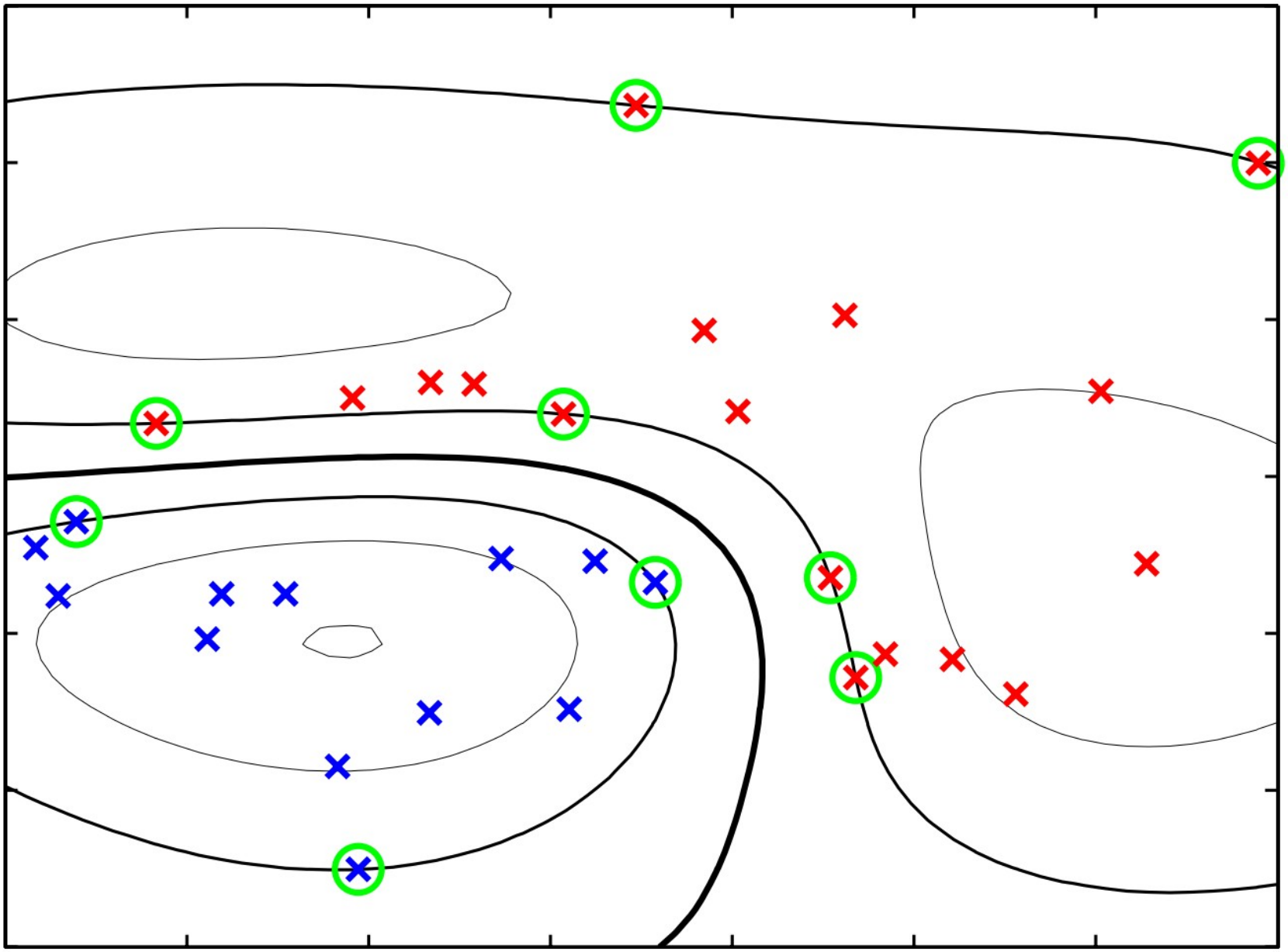
SVM: margin maximization

- **Margin**: distance between the blue lines



SVM – **non** separable case

SVM (Gaussian Kernel)



Key concepts

- **SVM**: maximum **margin**, **Lagrange** multip., **slack** variables..
  SVR: same idea, **regression** !
  + Ask ***Isabelle Guyon*** ;) *“Pattern recognition system using **support vectors**”, 1997*

References

- PCA: Bishop book, sec. 12.1.1, page 561-563
- Feature Maps:
  - a very clear explanation for ***feature maps***:
  https://scikit-learn.org/stable/modules/linear_model.html#polynomial-regression
  - *Bishop book*, sec. 6.1, p. 291-294
  - about ***approximate kernels***:https://scikit-learn.org/stable/modules/kernel_approximation.html
- **Algebra** reminder: *Bishop,* appendix C, p. 695-701 (only 6 pages !! 😀)
- **Regularization**: *Bishop*, sec. 3.1.4, p. 144-146
  See also Sec. 5.5, p. 256-271, for much much more (Neural Nets).
- Classification **Metrics** - a good summary:
  https://en.wikipedia.org/wiki/Confusion_matrix
- **SVMs**: https://see.stanford.edu/materials/aimlcs229/cs229-notes3.pdf (Andrew Ng's course, Stanford) – also: Bishop, sec. 7.1 (not all of it), p. 325-334

If you are **lost on everything** in this course ?
- ask more questions !
- read *Bishop*, Sec. 1.1, p. 1-11 (very simple intro)

References (FR)

- Find it for free: Aurélien Géron – *Hands-On Machine Learning with Scikit-Learn & TensorFlow*
- In French: almost the same table of contents – Aurélien Géron - *Machine Learning avec Scikit Learn - Mise en œuvre et cas concrets*

  I give the page numbers for the French edition, but same sections numbering in the English edition.
- **Train/Val/test**: sec 1.6, p.30
- **Métriques**: sec. 3.3, p. 86
- **Linear regression**: sec. 4.1, p. 108-113
- **Gradient Descent**: sec. 4.2, p. 113-122
- **Regularization**: sec. 4.5, p. 128
- **SVM**: sec 5.1, p. 145-148 (intuition sur les SVM)
  sec. 5.4 (under the hood) p. 156-… (jusqu'au Dual) : pour comprendre les maths sous-jacentes.
- **PCA**: sec. 8.1, p. 209-215 (read also the start of sec. 8, it's a good general intro to dimensional reduction)
- Perceptron, Kernel: only shown in part II (not present in French version)

Additional stuff

- List of standard kernels:
  https://scikit-learn.org/stable/modules/metrics.html#metrics
- `lecture5-SVM-showcase+exercises.ipynb`
- `lecture5-SVM-showcase+corrections.ipynb`
-