CS 66: Machine Learning

Prof. Sara Mathieson

Spring 2019

Lab 6: in-lab notes

Soft-margin SVM

non-separable case

regularization

$$\min_{\vec{w}, b, \vec{\xi}} \quad \frac{1}{2}\|\vec{w}\|^2 + C \cdot \sum_{i=1}^{n} \xi_i$$

s.t. $\underbrace{y_i(\vec{w}\cdot\vec{x}_i + b)}_{\text{functional margin.}} \geq 1 - \xi_i$

$\xi_i \geq 0$

C — hyper parameter to tune

new optimization problem

$\hat{y} = 1$

γ

"$-\xi_i$" 2.5

"$1 - \xi_i$" 0.2
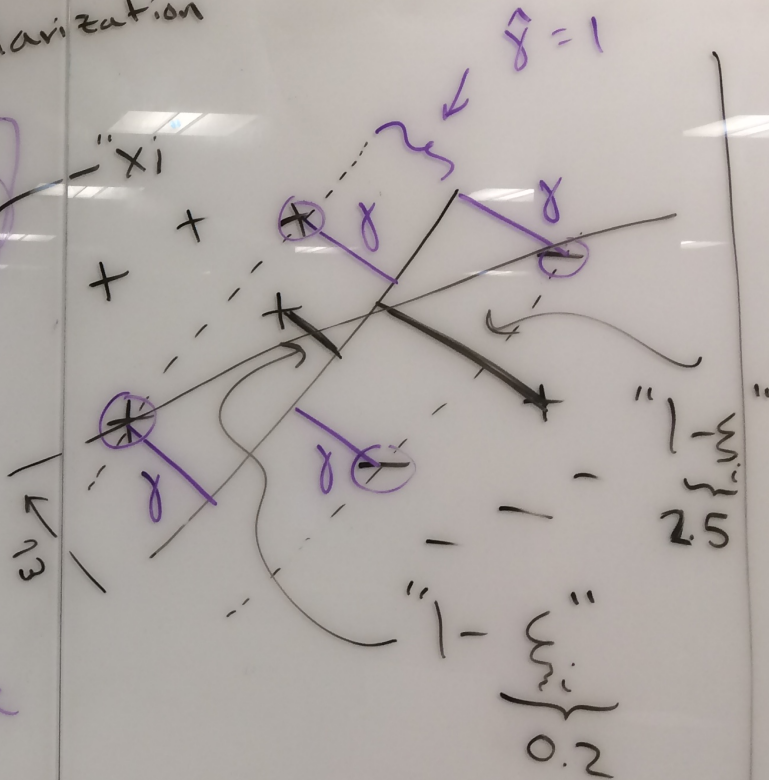
Optimization

Idea:

Coordinate

Ascent

Algorithm

$$\max_{\alpha} W(\alpha_1, \alpha_2 \cdots \alpha_n)$$
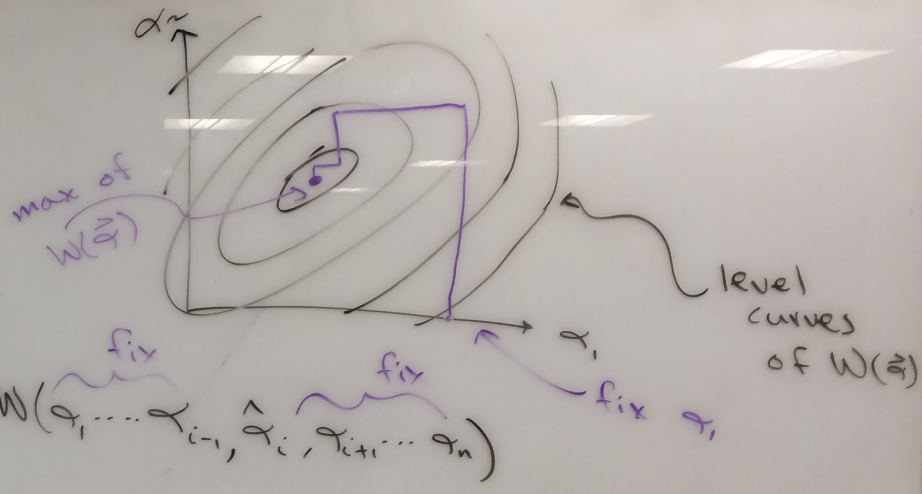
loop until convergence {

for $i = 1, \cdots n$ {

$\alpha_i \leftarrow \underset{\alpha_i}{\text{argmax}} \ W(\alpha_1, \cdots \alpha_{i-1}, \hat{\alpha}_i, \alpha_{i+1} \cdots \alpha_n)$

}

}

max of $W(\vec{\alpha})$

fix

fix

fix α_1
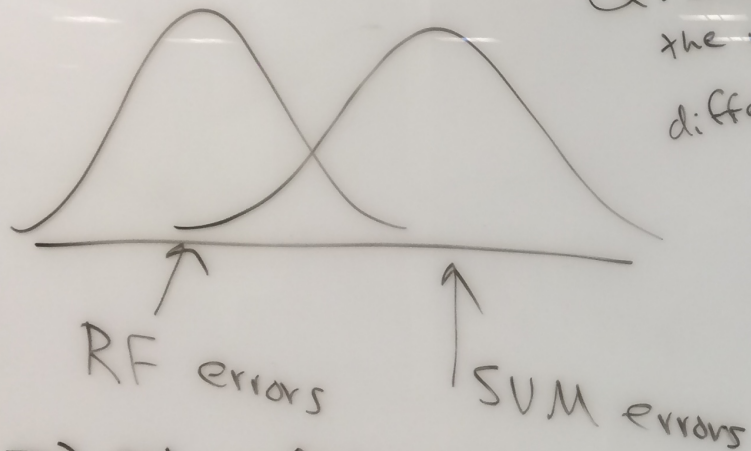
level curves of $W(\vec{\alpha})$

Compare algorithms statistically

Q: are the means different?

RF errors

SUM errors

=> get a distribution

=> use cross-validation.

Now

① non-separable case

② finding a's

③ goals of lab

Lab 6

Coding

* experiment 1
* experiment 2

Problem Set

* front (more proofs)
* back (numerical example)