**Experiment 1:**

Loss: hinge

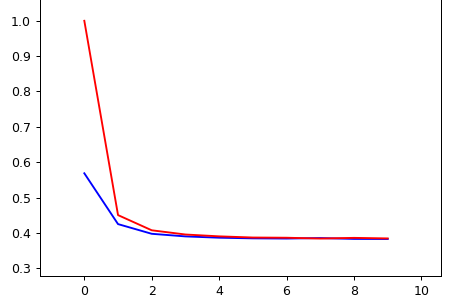
Cost:

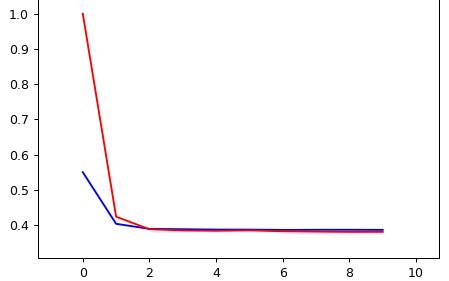
F(x) regular classification accuracy = 82%

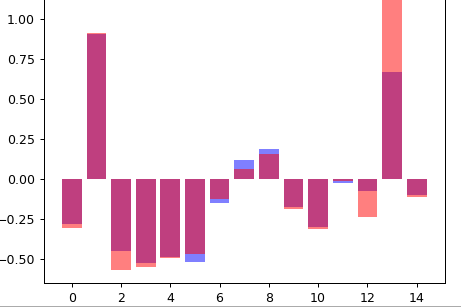
F(delta(x)) test accuracy = 69%

F’(delta(x)) test accuracy = 82%

F’(x) test accuracy = 65%

 F’

F

F’ parameters (blue) vs F parameters (red)

**Experiment 2:**

Loss: hinge

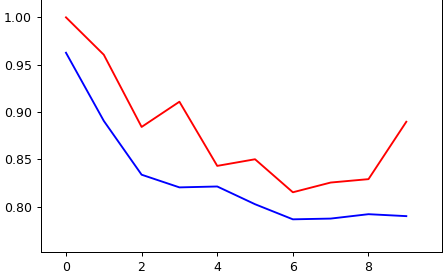
Cost:

F(x) regular classification accuracy = 82%

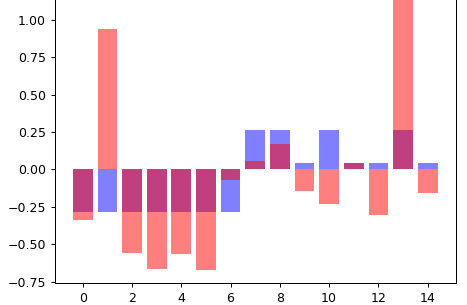
F(delta(x)) test accuracy = 52%

F’(delta(x)) test accuracy = 63%

F’(x) test accuracy = 52%

F’

F



**Experiment 3:**

Loss: hinge

Cost:

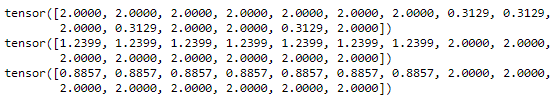
F(x) regular classification accuracy = 82%

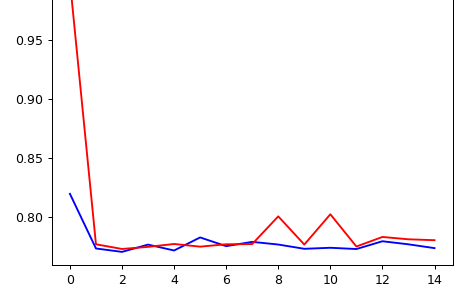
F(delta(x)) test accuracy = 49%

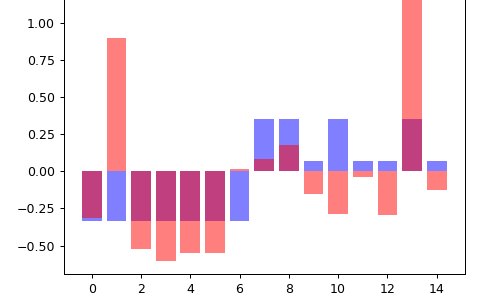
F’(delta(x)) test accuracy = 64%

F’(x) test accuracy = 63%

Weird X’ values:



F’



**Experiment 4:**

Loss: hinge

Cost:

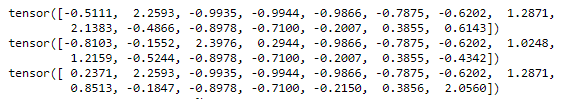
F(x) regular classification accuracy = 84%

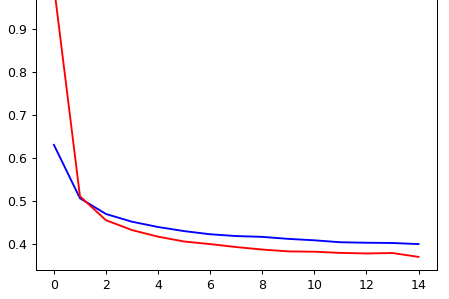
F(delta(x)) test accuracy = 57%

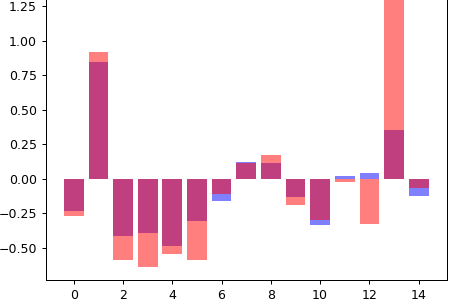
F’(delta(x)) test accuracy = 83%

F’(x) test accuracy = 49%

Weird X’ values:



F’



**Experiment 5:**

Loss: hinge

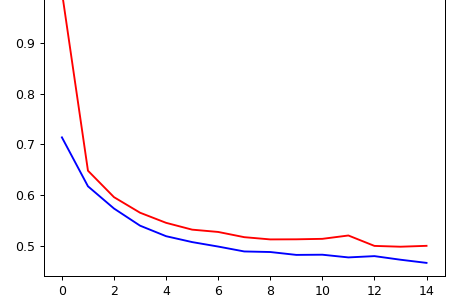
Cost:

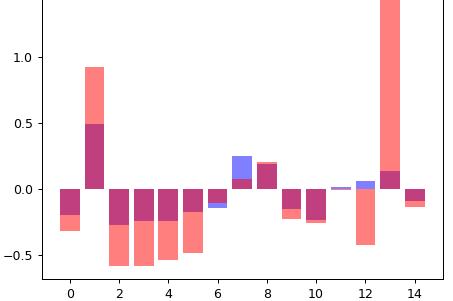
F(x) regular classification accuracy = 82%

F(delta(x)) test accuracy = 54%

F’(delta(x)) test accuracy = 79%

F’(x) test accuracy = 50%

F’



**Non-linearity (and x moves only when f(x) = -1):**

Tried

none of them were dpp (a rule-set for fast solving parametrized problems)

<https://www.cvxpy.org/tutorial/advanced/index.html#dpp>

**Ideas for next step:**

1. Regression: Train with loss:

*Red f is trained in a non-gaming environment*

1. Classification: Train with loss:

Findings:

With cost =

**Experiment 6: on synthetic data**

Loss: hinge

Cost:

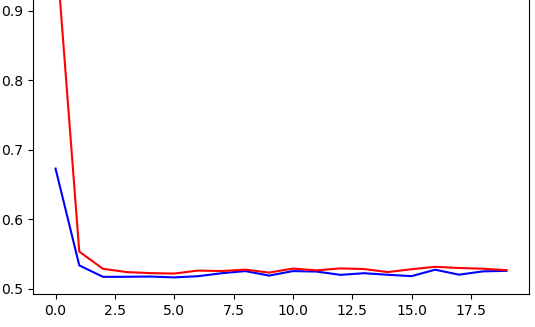
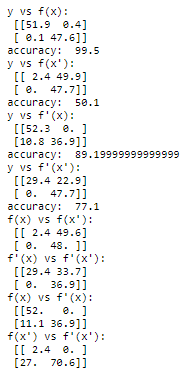
Gain:

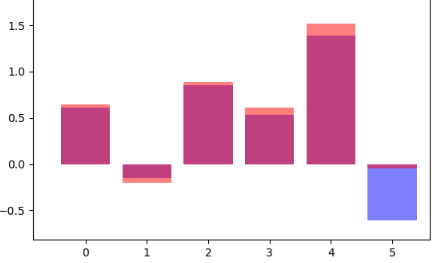
F(x) regular classification accuracy = 99%

F(delta(x)) test accuracy = 50%

F’(delta(x)) test accuracy = 89%

F’(x) test accuracy = 77%

F’



**Experiment 7:**

Loss: hinge

Cost:

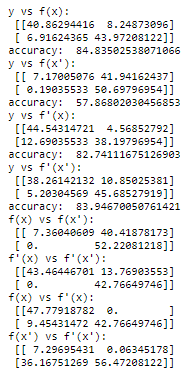
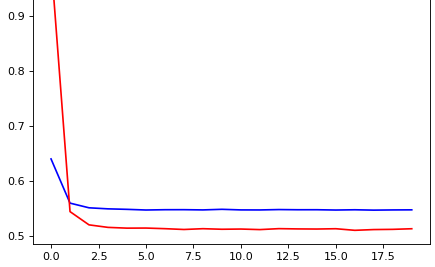
Gain:

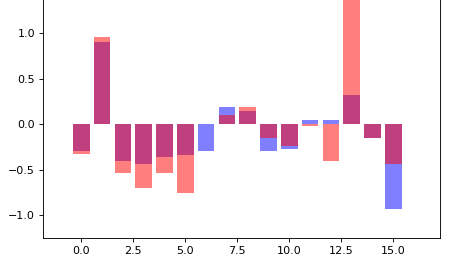
F(x) regular classification accuracy = 85%

F(delta(x)) test accuracy = 58%

F’(delta(x)) test accuracy = 84%

F’(x) test accuracy = 83%

 F’



Findings:

X’s move in the direction of w. each one with a unique difference