

## Follow the Regularized Leader

*Lecturer: Kris Kitani*

*Scribes: Grace Brueggman, Yuqing Qin*

# 1 Review

## 1.1 Follow the Leader (FTL)

In the previous lecture, we introduced the online convex optimization problem as a generalized version of online learning and the "Follow the Leader" (FTL) algorithm. FTL algorithm is a generic online optimization algorithm that follows two main steps. The first step is to estimate the parameters  $w$  such that it could minimize the cumulative loss up to time  $t-1$ . After the prediction step, it is going to receive the new loss function that will be used in the next timestamp. The algorithm is summarized below in Algorithm 1.

---

### Algorithm 1 Follow the Leader

---

1: <b>for</b> $t = 1, 2, \dots, T$ <b>do</b>	
2: $w^{(t)} = \arg \min_{w \in W} \sum_{i=1}^{t-1} f^{(i)}(w)$	▷ Parameter estimation
3:   RECEIVE $(f^{(t)} : W \rightarrow \mathbb{R})$	▷ Update loss function
4: <b>end for</b>	

---

We also saw two examples of FTL algorithm by applying different loss functions. With simple linear loss function as defined below, we saw it failed with the regret growing linearly.

$$f^{(t)}(w) = wz^{(z)}$$

The total loss of the FTL learner using the linear loss function is linearly to the time, and as a result, the regret also grows linearly.

$$L = \sum_{t=1}^T f^{(t)} = \sum_{t=2}^T 1 = T - 1$$

$$\text{Regret}(u) = T - 1 = O(T)$$

To further improve the performance, we also define a quadratic loss function, which in turn gives us lower regret bound. The quadratic loss is defined as below:

$$f^{(t)}(w) = \frac{1}{2} \|w - z^{(t)}\|_2^2$$

The regret bound for FTL with above quadratic loss function is better than the linear loss function's, which gives us a no regret algorithm.

$$Regret \leq 4L^2(\log(T) + 1) = O(\log T)$$

In today's lecture, we will show the proof of this better regret bound given by FTL with quadratic loss function and try to modify the linear loss function to make it perform better.

## 2 Summary

### 2.1 FTL with Quadratic Loss

The quadratic loss is defined as below:

$$f^{(t)}(w) = \frac{1}{2} \|w - z^{(t)}\|_2^2$$

We also assume

$$\|z\|_2^2 \leq L$$

---

**Algorithm 2** Follow the Leader with Quadratic Loss

---

1:	for $t = 1, 2, \dots, T$ do	
2:	$w^{(t)} = \arg \min_{w \in W} \sum_{i=1}^{t-1} f^{(i)}(w)$	▷ Parameter estimation
3:	RECEIVE $(f^{(t)}(w) = \frac{1}{2} \ w - z^{(t)}\ _2^2)$	▷ Quadratic loss function
4:	end for	

---

From the quadratic loss definition, we could minimize the loss by taking the derivatives of it and obtain the optimal  $w$ . The minimizer of a sum of squared difference is the average over the elements. As a result, we could rewrite line 2 in above algorithm to be:

$$w^{(t)} = \frac{1}{t-1} \sum_{i=1}^{t-1} z^{(i)}$$

#### 2.1.1 Regret Bound

The regret bound for FTL with quadratic loss is

$$Regret \leq 4L^2(\log(T) + 1)$$

To derive this regret bound, we use a new proof technique with only two steps.

- Upper-bound the regret using the one-step look ahead cheater
- Approximations and Algebra

By definition, the cumulative regret is defined as

$$R(u) = \sum_t [f^{(t)}(w^{(t)}) - f^{(t)}(u)]$$

The first term in this regret equation represents the loss of the online learner, and the second term is the loss of some hypothesis  $u$

The best hypothesis at timestamp  $t$  is actually the one that could minimize the cumulative loss up to time  $t$ . Therefore,

$$BestHypothesis = \arg \min_w \sum_{i=1}^t f^{(i)} = w^{(t+1)}$$

The best hypothesis at time  $t$  is the next time step's parameter  $w^{(t+1)}$ . We called this "one step look ahead cheater". With this, we could derive the regret bound as below:

$$R(u) = \sum_t [f^{(t)}(w^{(t)}) - f^{(t)}(u)] \leq \sum_t [f^{(t)}(w^{(t)}) - f^{(t)}(w^{(t+1)})]$$

To prove this regret bound, we only need to prove:

$$\sum_{t=1}^T [f^{(t)}(w^{(t+1)})] \leq \sum_{t=1}^T [f^{(t)}(u)]$$

Proof by Induction is to assume claim for  $T-1$  and prove it holds for  $T$ . We will use this strategy to prove the above. Therefore, we take one step before, and assume

$$\sum_{t=1}^{T-1} [f^{(t)}(w^{(t+1)})] \leq \sum_{t=1}^{T-1} [f^{(t)}(u)]$$

When we add  $f^{(T)}(w^{(T+1)})$  at time step  $T$  to both sides, we could get

$$\sum_{t=1}^T [f^{(t)}(w^{(t+1)})] \leq f^{(T)}(w^{(T+1)}) + \sum_{t=1}^{T-1} [f^{(t)}(u)]$$

The left hand side gets absorb the time step  $T$ 's loss, but the right hand side could not simply absorb that new loss term. Since the inequality holds for all values of  $u$ , and we can make the right hand side to be simpler by making the  $u = w^{(T+1)}$ . Therefore,

$$\sum_{t=1}^T [f^{(t)}(w^{(t+1)})] \leq \sum_{t=1}^T [f^{(t)}(w^{(T+1)})]$$

The only difference between the left hand side and the right hand side is on the  $w^{(T+1)}$ , the left hand side means we always use one step look ahead cheater (difference cheaters at each time step), the right hand side means we only use one single cheater (the cheater at the end of the sequence) for all time steps.

The minimizer by definition,

$$w^{(T+1)} = \arg \min_u \sum_{t=1}^T f^{(t)}(u) \leq u$$

Therefore, the above cumulative loss inequality's RHS can be upper bounded by all  $u$ , which is

$$\sum_{t=1}^T [f^{(t)}(w^{(t+1)})] \leq \sum_{t=1}^T [f^{(t)}(u)]$$

We have successfully proved the regret inequality without making any assumption yet. The regret bound is

$$R(u) = \sum_t [f^{(t)}(w^{(t)}) - f^{(t)}(u)] \leq \sum_t [f^{(t)}(w^{(t)}) - f^{(t)}(w^{(t+1)})]$$

Now, we will assume the loss function is quadratic, insert the loss into the RHS of the bound,

$$f^{(t)}(w^{(t)}) - f^{(t)}(w^{(t+1)}) = \frac{1}{2} \|w^{(t)} - z^{(t)}\|_2^2 - \frac{1}{2} \|w^{(t+1)} - z^{(t)}\|_2^2$$

To further simply this, we could rewrite the weight update into an incremental form. Recall that the minimizer of quadratic loss function is just the average. Instead of taking the average over all the values for each time step, we could simply incrementally add the new value with the previous mean value and normalize it. Therefore,

$$w^{(t+1)} = (1 - \frac{1}{t})w^{(t)} + (\frac{1}{t})z^{(t)}$$

To plug it in the regret bound,

$$\begin{aligned} f^{(t)}(w^{(t)}) - f^{(t)}(w^{(t+1)}) &= \frac{1}{2} \|w^{(t)} - z^{(t)}\|_2^2 - \frac{1}{2} \|(1 - \frac{1}{t})w^{(t)} + (\frac{1}{t})z^{(t)} - z^{(t)}\|_2^2 \\ &= \frac{1}{2} (1 - (1 - \frac{1}{t})^2) \|w^{(t)} - z^{(t)}\|_2^2 \end{aligned}$$

In the graph at Figure 1 we depict two functions.

$$(1) f(t) = \frac{1}{t}$$

$$(2) f(t) = \frac{1}{2}(1 - (1 - \frac{1}{t})^2)$$

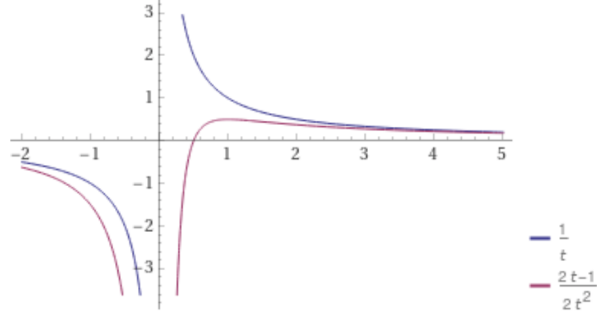


Figure 1: The known inequality

Because of the relationship between these functions, where we can create the following inequality:

$$\frac{1}{2}(1 - (1 - \frac{1}{t})^2) \leq \frac{1}{t}$$

With this inequality, we can derive the upper bound for the FTL regret.

$$\frac{1}{2}(1 - (1 - \frac{1}{t})^2) \|w^{(t)} - z^{(z)}\|_2^2 \leq \frac{1}{t} \|w^{(t)} - z^{(z)}\|_2^2$$

Recall that we made prior assumptions that

$$L = \max_t \|z^{(t)}\|$$

Since the average is always less or equal to the max,

$$\|w^{(t)}\| \leq L$$

From this, we could derive below inequality based on the triangle inequality. We could think of two vectors, each might have a magnitude of  $L$ , therefore,

$$\|w^{(t)} - z^{(t)}\| \leq 2L$$

As a result, the regret at time  $t$  is upper bounded by

$$\frac{1}{2}(1 - (1 - \frac{1}{t})^2) \|w^{(t)} - z^{(z)}\|_2^2 \leq \frac{1}{t} \|w^{(t)} - z^{(z)}\|_2^2 \leq (\frac{1}{t}) 4L^2$$

The regret bound over all time steps is

$$R(u) \leq \sum_t [f^{(t)}(w^{(t)}) - f^{(t)}(w^{(t+1)})] \leq \sum_t (\frac{1}{t}) 4L^2$$

Another useful inequality to simply this is

$$\sum_{t=1}^T \left(\frac{1}{t}\right) \leq \log(T) + 1$$

Plug it in and we can get the final regret bound,

$$R(u) \leq 4L^2(\log(T) + 1)$$

To summarize the FTL with quadratic loss function, we could see the algorithm is a no regret algorithm from the above regret bound. Also the weight update is via averaging, which updates the weights smoothly. The weight does not change dramatically over time and it ensure the stability for the algorithm.

## 2.2 Follow the Regularized Leader (FTRL)

The generic version of the Follow the Regularized Leader is just like FTL with a regularizer term added in the parameter estimation. We can even think of FTL as a special case of FTRL where  $\psi = 0$ . In Algorithm 3 below,  $S$  is a convex set.

---

### Algorithm 3 Follow the Regularized Leader

---

- 1: **for**  $t = 1, 2, \dots, T$  **do**
  - 2:    $w^{(t)} = \arg \min_w \sum_{i=1}^{t-1} f^{(i)}(w) + \psi(w)$  ▷ Parameter estimation
  - 3:   RECEIVE  $(f^{(t)} : S \rightarrow \mathbb{R})$  ▷ Update loss function
  - 4: **end for**
- 

### 2.2.1 Follow the Regularized Leader (FTRL) with Linear Loss and Quadratic Regularization

One way to structure FTRL is with quadratic regularization and linear loss. With quadratic regularization, we use the following to calculate  $\psi(w)$ , such that the regularizer is minimized when the parameters  $w = 0$  and it induces an exponential penalty when the parameters are large:

$$\psi(w) = \frac{1}{2\eta} \|w\|_2^2$$

To calculate a linear loss, we use the function:

$$f^{(t)} = w \cdot z^{(t)}$$

As with all losses, it is a more optimized online learner solution when we minimize this term incrementally.

---

**Algorithm 4** Follow the Regularized Leader with Quadratic Regularization and Linear Loss

---

```
1: for  $t = 1, 2, \dots, T$  do  
2:    $w^{(t)} = \arg \min_{w \in S} (\frac{1}{2\eta} \|w\|_2^2 + w \cdot \sum_{i=1}^{t-1} z^{(i)})$  ▷ Parameter estimation  
3:   RECEIVE ( $f^{(t)}(w) = w \cdot z^{(t)}$ ) ▷ Update loss function  
4: end for
```

---

To find the minimum of the parameter estimation step in Algorithm 4, we have to first take the derivative of the cumulative loss with respect to  $w$  and then solve for  $w$ .

$$\frac{d}{dw} (\frac{1}{2\eta} \|w\|_2^2 + w \cdot \sum_{i=1}^{t-1} z^{(i)}) = 0$$

$$\frac{1}{2\eta} 2w + \sum_{i=1}^{t-1} z^{(i)} = 0$$

$$w^{(t)} = -\eta \sum_{i=1}^{t-1} z^{(i)}$$

Another way to look at  $w$  is in a recursive form where the new parameter is in terms of the old parameter and the incremental update is proportional to the gradient of the loss function at  $t$ ,  $z^{(t)}$ , such that:

$$w^{(t)} = -\eta (\sum_{i=1}^{t-2} z^{(i)} + z^{(t-1)})$$

$$w^{(t)} = -\eta \sum_{i=1}^{t-2} z^{(i)} - \eta z^{(t-1)}$$

$$w^{(t)} = w^{(t-1)} - \eta z^{(t-1)}$$

### 2.2.2 Regret Bound for FTRL with a Euclidean Regularizer and Linear Loss

**Proof of Regret Bound:**

$$R^{(T)}(u) \leq [\psi(u) - \psi(w^{(1)})] + \sum_{t=1}^T [f^{(t)}(w^{(t)}) - f^{(t)}(w^{(t+1)})]$$

We then substitute in definitions. We also set  $\psi(w^{(1)}) = 0$ , since we know  $w^{(1)}$  is the argmin of the L2 norm regularizer, which we previously determined is minimized at 0.

$$\begin{aligned} &= [\frac{1}{2\eta} \|u\|_2^2 - 0] + \sum_{t=1}^T [\langle w^{(t)}, z^{(t)} \rangle - \langle w^{(t+1)}, z^{(t)} \rangle] \\ &= \frac{1}{2\eta} \|u\|_2^2 + \sum_{t=1}^T [\langle w^{(t)} - w^{(t+1)}, z^{(t)} \rangle] \end{aligned}$$

Given the recursive form for the update parameter step derived above in Section 2.2.1, we know:

$$w^{(t+1)} = w^{(t)} - \eta z^{(t)}$$

$$w^{(t)} - w^{(t+1)} = \eta z^{(t)}$$

which can now be substituted into the following:

$$R^{(T)}(u) \leq \frac{1}{2\eta} \|u\|_2^2 + \sum_{t=1}^T \eta \|z^{(t)}\|^2$$

Since this is the upper bound, we will establish the following:

$$L = \max_z \|z\|_2$$

$$B = \max_{u \in S} \|u\|_2$$

Which, when substituted into the inequality, we have the cleaner solution:

$$R^{(T)}(u) \leq \frac{1}{2\eta} B^2 + \eta T L^2$$

The optimal  $\eta$  can be determined by taking the derivative of the RHS with respect to  $\eta$  and solving for  $\eta$ .

$$\frac{d}{d\eta} \left( \frac{1}{2\eta} B^2 + \eta T L^2 \right) = 0$$

$$-\frac{1}{2\eta^2} B^2 + T L^2 = 0$$

$$\frac{2 T L^2}{B^2} = \frac{1}{\eta^2}$$

$$\frac{B^2}{2 T L^2} = \eta^2$$

$$\frac{B}{L \sqrt{2T}} = \eta$$

We then take this value of  $\eta$  and put it in the inequality to get the upper bound of the regret:

$$R^{(T)}(u) \leq \frac{1}{2 \left( \frac{B}{L \sqrt{2T}} \right)} B^2 + \left( \frac{B}{L \sqrt{2T}} \right) T L^2$$

$$R^{(T)}(u) \leq B L \sqrt{2T}$$

This means FLRL with a linear loss and quadratic regularization is a no regret algorithm, or more explicitly, over a long time, the average regret will go to zero.

In relation to Convex Optimization, any sequence of Lipschitz loss functions, not just linear, and any convex regularization, not just quadratic, will also achieve stabilized results.