FOLLOW THE REGULARIZED LEADER

Advanced Machine Learning

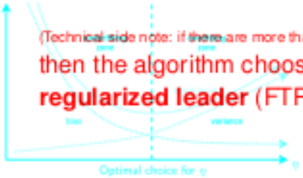
Follow the regularized leader

- To overcome the shortcomings of the FTL algorithm, one can incorporate a regularization function $\psi : \mathcal{A} \to \mathbb{R}_+$ into the action choice of FTL, which leads to more stability.
- To be more precise, let for $t \geq 1$

$$a_t^{\mathrm{FTRL}} \in \arg\min_{a \in \mathcal{A}} \left(\psi(a) + \sum\nolimits_{s=1}^{t-1} (a, z_s) \right),$$

(Technical side note: if there are more than one minimum, then one of them is chosen.)

then the algorithm choosing a_t^{FTRL} in time step t is called the **Follow the regularized leader** (FTRL) algorithm.



Learning goals

- Get to know FTRL as a stable alternative for FTL
- See a suitable regularization for OLO problems

FOLLOW THE REGULARIZED LEADER

- To overcome the shortcomings of the FTL algorithm, one can incorporate a regularization function $\psi : \mathcal{A} \to \mathbb{R}_+$ into the action choice of FTL, which leads to more stability.
- To be more precise, let for $t \geq 1$

$$a_t^{\mathrm{FTRL}} \in \arg\min_{a \in \mathcal{A}} \left(\psi(a) + \sum_{s=1}^{t-1} \ell(a, z_s) \right),$$

(Technical side note: if there are more than one minimum, then one of them is chosen.)

then the algorithm choosing a_t^{FTRL} in time step t is called the **Follow the regularized leader** (FTRL) algorithm.

- *Interpretation:* The algorithm predicts a_t as the element in $\mathcal{A}$, which minimizes the regularization function plus the cumulative loss so far over the previous $t-1$ time periods.

FOLLOW THE REGULARIZED LEADER

- To overcome the shortcomings of the FTL algorithm, one can incorporate a regularization function $\psi : \mathcal{A} \to \mathbb{R}_+$ into the action choice of FTL, which leads to more stability.
- To be more precise, let for $t \geq 1$

$$a_t^{\mathrm{FTRL}} \in \arg\min_{a \in \mathcal{A}} \left(\psi(a) + \sum_{s=1}^{t-1} \ell(a, z_s) \right),$$

(Technical side note: if there are more than one minimum, then one of them is chosen.)

then the algorithm choosing a_t^{FTRL} in time step t is called the **Follow the regularized leader** (FTRL) algorithm.

- *Interpretation:* The algorithm predicts a_t as the element in $\mathcal{A}$, which minimizes the regularization function plus the cumulative loss so far over the previous $t-1$ time periods.
- Obviously, the behavior of the FTRL algorithm is depending heavily on the choice of the regularization function ψ. If $\psi \equiv 0$, then FTRL equals FTL.

REGULARIZATION IN ONLINE LEARNING VS. BATCH LEARNING

- Note that in the batch learning scenario, the learner seeks to optimize an objective function which is the sum of the training loss and a regularization function:

$$\min_{\theta \in \mathbb{R}^p} \sum_{i=1}^{n} L(y^{(i)}, \theta) + \lambda \psi(\theta),$$

where $\lambda \geq 0$ is some regularization parameter.

- To overcome the shortcomings of the FTL algorithm, one can incorporate a regularization function $\psi : \mathcal{A} \to \mathbb{R}$ into the action choice of FTL, which leads to more stability.
- To be more precise, let for $t \geq 1$

$$a_t^{\text{FTRL}} \in \arg\min_{a \in \mathcal{A}} \left(\psi(a) + \sum_{s=1}^{t-1} l(a, z_s) \right),$$

(Technical side note: if there are more than one minimum, then one of them is chosen.)

then the algorithm choosing a_t^{FTRL} in time step t is called the **Follow the regularized leader** (FTRL) algorithm.

- *Interpretation:* The algorithm predicts a_t as the element in $\mathcal{A}$, which minimizes the regularization function plus the cumulative loss so far over the previous $t-1$ time periods.
- Obviously, the behavior of the FTRL algorithm is depending heavily on the choice of the regularization function ψ. If $\psi \equiv 0$, then FTRL equals FTL.

REGULARIZATION IN ONLINE LEARNING VS. BATCH LEARNING

- Note that in the batch learning scenario, the learner seeks to optimize an objective function which is the sum of the training loss and a regularization function:

$$\min_{\boldsymbol{\theta} \in \mathbb{R}^p} \sum_{i=1}^{n} L(y^{(i)}, \boldsymbol{\theta}) + \lambda \psi(\boldsymbol{\theta}),$$

  where $\lambda \geq 0$ is some regularization parameter.

- Here, the regularization function is part of the whole objective function, which the learner seeks to minimize.

REGULARIZATION IN ONLINE LEARNING VS. BATCH LEARNING

- Note that in the batch learning scenario, the learner seeks to optimize an objective function which is the sum of the training loss and a regularization function:

$$\min_{\boldsymbol{\theta} \in \mathbb{R}^p} \sum_{i=1}^{n} L(y^{(i)}, \boldsymbol{\theta}) + \lambda \psi(\boldsymbol{\theta}),$$

  where $\lambda \geq 0$ is some regularization parameter.

- Here, the regularization function is part of the whole objective function, which the learner seeks to minimize.

- However, in the online learning scenario the regularization function does (usually) not appear in the regret the learner seeks to minimize, but the regularization function is only part of the action/decision rule at each time step.

REGRET ANALYSIS OF FTRL: A HELPFUL LEMMA

BATCH LEARNING

- **Lemma:** Let $a_1^{\mathrm{FTRL}}, a_2^{\mathrm{FTRL}}, \ldots$ be the sequence of actions coming used by the FTRL algorithm for the environmental data sequence $z_1, z_2, \ldots$. Then, for all $\tilde{a} \in \mathcal{A}$ we have

$$R_T^{\mathrm{FTRL}}(\tilde{a}) = \sum_{t=1}^{T} \left((a_t^{\mathrm{FTRL}}, z_t) - (\tilde{a}, z_t) \right)$$

$$\leq \psi(\tilde{a}) - \psi(a_1^{\mathrm{FTRL}}) + \sum_{t=1}^{T} \left((a_t^{\mathrm{FTRL}}, z_t) - (a_{t+1}^{\mathrm{FTRL}}, z_t) \right).$$

- Note that in the batch learning scenario, the learner seeks to optimize an objective function which is the sum of the training loss and a regularization function:

$$\min_{\theta \in \mathbb{R}^p} \sum_{i=1} L(y^{(i)}, \theta) + \lambda \psi(\theta),$$

where $\lambda \geq 0$ is some regularization parameter.

- Here, the regularization function is part of the whole objective function, which the learner seeks to minimize.

- However, in the online learning scenario the regularization function does (usually) not appear in the regret the learner seeks to minimize, but the regularization function is only part of the action/decision rule at each time step.

REGRET ANALYSIS OF FTRL: A HELPFUL LEMMA

- **Lemma:** Let $a_1^{\text{FTRL}}, a_2^{\text{FTRL}}, \ldots$ be the sequence of actions coming used by the FTRL algorithm for the environmental data sequence $z_1, z_2, \ldots$. Then, for all $\tilde{a} \in \mathcal{A}$ we have

$$R_T^{\text{FTRL}}(\tilde{a}) = \sum_{t=1}^{T} \left(\ell(a_t^{\text{FTRL}}, z_t) - \ell(\tilde{a}, z_t) \right)$$

$$\leq \psi(\tilde{a}) - \psi(a_1^{\text{FTRL}}) + \sum_{t=1}^{T} \left(\ell(a_t^{\text{FTRL}}, z_t) - \ell(a_{t+1}^{\text{FTRL}}, z_t) \right).$$

- *Interpretation*: the regret of the FTRL algorithm is bounded by the difference of cumulated losses of itself compared to its one-step lookahead cheater version and an additional regularization difference term.

$\Rightarrow$ We have seen an analogous result for FTL!

(The proof is similar.)

FTRL FOR ONLINE LINEAR OPTIMIZATION

- In the following, we analyze the FTRL algorithm for the linear loss $(a, z) = a^\top z$ for online linear optimization (OLO) problems.
- For this purpose, the squared L2-norm regularization will be used:

$$\psi(a) = \frac{1}{2\eta}\|a\|_2^2 = \frac{a^\top a}{2\eta},$$

where η is some positive scalar, the *regularization magnitude.*

REGRET ANALYSIS OF FTRL: LEMMA

- Lemma: Let $a_1^{\mathrm{FTRL}}, a_2^{\mathrm{FTRL}}, \ldots$ be the sequence of actions being used by the FTRL algorithm for the environmental data sequence $z_1, z_2, \ldots$. Then, for all $\bar{a} \in \mathcal{A}$ we have

$$R_T^{\mathrm{FTRL}}(\bar{a}) = \sum_{t=1}^{T} \left(L(a_t^{\mathrm{FTRL}}, z_t) - L(\bar{a}, z_t) \right)$$

$$\leq \psi(\bar{a}) - \psi(a_1^{\mathrm{FTRL}}) + \sum_{t=1}^{T} \left(L(a_t^{\mathrm{FTRL}}, z_t) - L(a_{t+1}^{\mathrm{FTRL}}, z_t) \right).$$

- *Interpretation*: the regret of the FTRL algorithm is bounded by the difference of cumulated losses of itself compared to its one-step lookahead cheater version and an additional regularization difference term.

$\Rightarrow$ We have seen an analogous result for FTL!

(The proof is similar.)

FTRL FOR ONLINE LINEAR OPTIMIZATION

- In the following, we analyze the FTRL algorithm for the linear loss $\ell(a, z) = a^\top z$ for online linear optimization (OLO) problems.
- For this purpose, the squared L2-norm regularization will be used:

$$\psi(a) = \frac{1}{2\eta} ||a||_2^2 = \frac{a^\top a}{2\eta},$$

where η is some positive scalar, the *regularization magnitude*.

- It is straightforward to compute that if $\mathcal{A} = \mathbb{R}^d$, then

$$a_t^{\mathrm{FTRL}} = -\eta \sum_{s=1}^{t-1} z_s.$$

FTRL FOR ONLINE LINEAR OPTIMIZATION

- In the following, we analyze the FTRL algorithm for the linear loss $\ell(a, z) = a^\top z$ for online linear optimization (OLO) problems.
- For this purpose, the squared L2-norm regularization will be used:

$$\psi(a) = \frac{1}{2\eta} \|a\|_2^2 = \frac{a^\top a}{2\eta},$$

where η is some positive scalar, the *regularization magnitude*.

- It is straightforward to compute that if $\mathcal{A} = \mathbb{R}^d$, then

$$a_t^{\mathrm{FTRL}} = -\eta \sum_{s=1}^{t-1} z_s.$$

- Hence, in this case we have for the FTRL algorithm the following update rule

$$a_{t+1}^{\mathrm{FTRL}} = a_t^{\mathrm{FTRL}} - \eta\, z_t, \qquad t = 1, \ldots, T-1.$$

FTRL FOR ONLINE LINEAR OPTIMIZATION

- In the following, we analyze the FTRL algorithm for the linear loss $\ell(a, z) = a^\top z$ for online linear optimization (OLO) problems.
- For this purpose, the squared L2-norm regularization will be used:

$$\psi(a) = \frac{1}{2\eta} ||a||_2^2 = \frac{a^\top a}{2\eta},$$

where η is some positive scalar, the *regularization magnitude*.

- It is straightforward to compute that if $\mathcal{A} = \mathbb{R}^d$, then

$$a_t^{\mathrm{FTRL}} = -\eta \sum_{s=1}^{t-1} z_s.$$

- Hence, in this case we have for the FTRL algorithm the following update rule

$$a_{t+1}^{\mathrm{FTRL}} = a_t^{\mathrm{FTRL}} - \eta\, z_t, \qquad t = 1, \ldots, T-1.$$

Interpretation: $-z_t$ is the *direction* in which the update of a_t^{FTRL} to a_{t+1}^{FTRL} is conducted with *step size* η in order to reduce the loss.

FTRL FOR OLO: THEORETICAL GUARANTEES

- **Proposition:** Using the FTRL algorithm with the squared L2-norm regularization on any online linear optimization (OLO) problem with $\mathcal{A} \subseteq \mathbb{R}^d$ leads to a regret of FTRL with respect to any action $\tilde{a} \in \mathcal{A}$ of

$$R_T^{FTRL}(\tilde{a}) \leq \frac{1}{2\eta}\|\tilde{a}\|_2^2 + \eta \sum_{t=1}^{T} \|z_t\|_2^2 .$$

- In the following, we analyze the FTRL algorithm for the linear loss $l(a, z) = a^\top z$ for online linear optimization (OLO) problems.
- For this purpose, the squared L2-norm regularization will be used:

$$\psi(a) = \frac{1}{2\eta}\|a\|_2^2 = \frac{a^\top a}{2\eta},$$

where η is some positive scalar, the *regularization magnitude*.

- It is straightforward to compute that if $\mathcal{A} = \mathbb{R}^d$, then

$$a_t^{\text{FTRL}} = -\eta \sum_{s=1}^{t-1} z_s.$$

- Hence, in this case we have for the FTRL algorithm the following update rule

$$a_{t+1}^{\text{FTRL}} = a_t^{\text{FTRL}} - \eta\, z_t, \qquad t = 1, \ldots, T-1.$$

Interpretation: $-z_t$ is the *direction* in which the update of a_t^{FTRL} to a_{t+1}^{FTRL} is conducted with *step size* η in order to reduce the loss.

FTRL FOR OLO: THEORETICAL GUARANTEES

- **Proposition:** Using the FTRL algorithm with the squared L2-norm regularization on any online linear optimization (OLO) problem with $\mathcal{A} \subset \mathbb{R}^d$ leads to a regret of FTRL with respect to any action $\tilde{a} \in \mathcal{A}$ of

$$R_T^{FTRL}(\tilde{a}) \leq \frac{1}{2\eta} \|\tilde{a}\|_2^2 + \eta \sum_{t=1}^{T} \|z_t\|_2^2 .$$

- We will show the result only for the case $\mathcal{A} = \mathbb{R}^d$.
- For the more general case, where $\mathcal{A}$ is a strict subset of $\mathbb{R}^d$, we need a slight modification of the update formula above:

$$a_t^{\text{FTRL}} = \Pi_{\mathcal{A}}\left(-\eta \sum_{i=1}^{t-1} z_i\right) = \arg\min_{a \in \mathcal{A}} \left\| a - \eta \sum_{i=1}^{t-1} z_i \right\|_2^2 .$$

In words, the action of the FTRL algorithm has to be projected onto the set $\mathcal{A}$. Here, $\Pi_{\mathcal{A}} : \mathbb{R}^d \to \mathcal{A}$ is the projection onto $\mathcal{A}$.

(The proof is essentially the same, except that the Cauchy-Schwarz inequality is used in between.)

FTRL FOR OLO: THEORETICAL GUARANTEES

- **Proposition:** Using the FTRL algorithm with the squared L2-norm regularization on any online linear optimization (OLO) problem with $\mathcal{A} \subset \mathbb{R}^d$ leads to a regret of FTRL with respect to any action $\tilde{a} \in \mathcal{A}$ of

$$R_T^{\mathrm{FTRL}}(\tilde{a}) \leq \frac{1}{2\eta}\|\tilde{a}\|_2^2 + \eta \sum_{t=1}^{T} \|z_t\|_2^2 .$$

- **Proof:**

Reminder (1): $R_T^{\mathrm{FTRL}}(\tilde{a}) \leq \psi(\tilde{a}) - \psi(a_1^{\mathrm{FTRL}}) + \sum_{t=1}^{T} \left(\langle a_t^{\mathrm{FTRL}}, z_t \rangle - \langle a_{t+1}^{\mathrm{FTRL}}, z_t \rangle \right).$

Reminder (2): $a_{t+1}^{\mathrm{FTRL}} = a_t^{\mathrm{FTRL}} - \eta z_t, \quad t = 1, \dots, T-1.$

- For sake of brevity, we write $a_1, a_2, \dots$ for $a_1^{\mathrm{FTRL}}, a_2^{\mathrm{FTRL}}, \dots$

- We will show the result only for the case $\mathcal{A} = \mathbb{R}^d$.

- For the more general case, where $\mathcal{A}$ is a strict subset of $\mathbb{R}^d$, we need a slight modification of the update formula above:

$$a_t^{\mathrm{FTRL}} = \Pi_{\mathcal{A}}\left(-\eta \sum_{i=1}^{t-1} z_i\right) = \arg\min_{a \in \mathcal{A}} \left\| a - \eta \sum_{i=1}^{t-1} z_i \right\|_2^2 .$$

In words, the action of the FTRL algorithm has to be projected onto the set $\mathcal{A}$. Here, $\Pi_{\mathcal{A}} : \mathbb{R}^d \to \mathcal{A}$ is the projection onto $\mathcal{A}$.

(The proof is essentially the same, except that the Cauchy-Schwarz inequality is used in between.)

FTRL FOR OLO: THEORETICAL GUARANTEES

- **Proof:**

  Reminder (1): $R_T^{FTRL}(\tilde{a}) \le \psi(\tilde{a}) - \psi(a_1^{FTRL}) + \sum_{t=1}^{T} \left(\langle a_t^{FTRL}, z_t \rangle - \langle a_{t+1}^{FTRL}, z_t \rangle \right).$

  Reminder (2): $a_{t+1}^{FTRL} = a_t^{FTRL} - \eta z_t, \quad t = 1, \dots, T-1.$

- For sake of brevity, we write $a_1, a_2, \dots$ for $a_1^{FTRL}, a_2^{FTRL}, \dots$
- With this,

$$
\begin{aligned}
R_T^{FTRL}(\tilde{a}) &\le \psi(\tilde{a}) - \psi(a_1) + \sum_{t=1}^{T} (\langle a_t, z_t \rangle - \langle a_{t+1}, z_t \rangle) && \text{(Reminder (1))} \\
&\le \frac{1}{2\eta} \|\tilde{a}\|_2^2 + \sum_{t=1}^{T} (a_t^\top z_t - a_{t+1}^\top z_t) && (\psi(a_1) \ge 0 \text{ and definition of } \psi) \\
&= \frac{1}{2\eta} \|\tilde{a}\|_2^2 + \sum_{t=1}^{T} (a_t^\top - a_{t+1}^\top) z_t && \text{(Distributivity)} \\
&= \frac{1}{2\eta} \|\tilde{a}\|_2^2 + \eta \sum_{t=1}^{T} \|z_t\|_2^2. && \text{(Reminder (2))}
\end{aligned}
$$

□

FTRL FOR OLO: THEORETICAL GUARANTEES

- Proof:
- Interpretation of the terms in the proposition, i.e., of

$$R_T^{FTRL}(\bar{a}) \leq \frac{1}{2\eta}||\bar{a}||_2^2 + \eta \sum_{t=1}^{T} ||z_t||_2^2 \,;$$

Reminder (1): $R_T^{\mathrm{FTRL}}(\bar{a}) \leq \psi(\bar{a}) - \psi(a_1^{\mathrm{FTRL}}) + \sum_{t=1}^{T} \left(L(a_t^{\mathrm{FTRL}}, z_t) - L(a_{t+1}^{\mathrm{FTRL}}, z_t) \right).$

Reminder (2): $a_{t+1}^{\mathrm{FTRL}} = a_t^{\mathrm{FTRL}} - \eta\, z_t, \qquad t = 1, \ldots, T-1.$

- For sake of brevity, we write $a_1, a_2, \ldots$ for $a_1^{\mathrm{FTRL}}, a_2^{\mathrm{FTRL}}, \ldots$
- With this,

$$\begin{aligned}
R_T^{FTRL}(\bar{a}) &\leq \psi(\bar{a}) - \psi(a_1) + \sum\nolimits_{t=1}^{T} (L(a_t, z_t) - L(a_{t+1}, z_t)) && \text{(Reminder (1))} \\
&\leq \frac{1}{2\eta} ||\bar{a}||_2^2 + \sum\nolimits_{t=1}^{T} (a_t^\top z_t - a_{t+1}^\top z_t) && (\psi(a_1) \geq 0 \text{ and definition of } \psi) \\
&= \frac{1}{2\eta} ||\bar{a}||_2^2 + \sum\nolimits_{t=1}^{T} (a_t^\top - a_{t+1}^\top) z_t && \text{(Distributivity)} \\
&= \frac{1}{2\eta} ||\bar{a}||_2^2 + \eta \sum\nolimits_{t=1}^{T} ||z_t||_2^2 . && \text{(Reminder (2))}
\end{aligned}$$

□

FTRL FOR OLO: THEORETICAL GUARANTEES

- Interpretation of the terms in the proposition, i.e., of

$$R_T^{FTRL}(\tilde{a}) \leq \frac{1}{2\eta}||\tilde{a}||_2^2 + \eta\sum_{t=1}^{T}||z_t||_2^2 :$$

  - $||\tilde{a}||_2^2$ represents a *bias term:* The regret upper bound of FTRL is always biased by the term $||\tilde{a}||_2^2$. The impact of the bias term can be reduced by a higher regularization magnitude, i.e., a higher choice of η.

FTRL FOR OLO: THEORETICAL GUARANTEES

- Interpretation of the terms in the proposition, i.e., of

$$R_T^{FTRL}(\tilde{a}) \leq \frac{1}{2\eta}||\tilde{a}||_2^2 + \eta\sum_{t=1}^{T}||z_t||_2^2 :$$

  - $||\tilde{a}||_2^2$ represents a *bias term:* The regret upper bound of FTRL is always biased by the term $||\tilde{a}||_2^2$. The impact of the bias term can be reduced by a higher regularization magnitude, i.e., a higher choice of η.
  - $\sum_{t=1}^{T}||z_t||_2^2$ represents a *"variance" term*: The more the environment data z_t varies, the larger this term. Hence, for a high variance a smaller regularization magnitude is needed, i.e., a smaller choice of η.

FTRL FOR OLO: THEORETICAL GUARANTEES

- Interpretation of the terms in the proposition, i.e., of

$$R_T^{FTRL}(\tilde{a}) \leq \frac{1}{2\eta}||\tilde{a}||_2^2 + \eta \sum_{t=1}^{T} ||z_t||_2^2 :$$

  - $||\tilde{a}||_2^2$ represents a *bias term:* The regret upper bound of FTRL is always biased by the term $||\tilde{a}||_2^2$. The impact of the bias term can be reduced by a higher regularization magnitude, i.e., a higher choice of η.
  - $\sum_{t=1}^{T} ||z_t||_2^2$ represents a *"variance" term*: The more the environment data z_t varies, the larger this term. Hence, for a high variance a smaller regularization magnitude is needed, i.e., a smaller choice of η.

- Thus, we have a trade-off for the optimal choice of η : Making η large, leads to a smaller bias but at the expense of a higher variance and making η small leads to a smaller variance at the expense of a higher bias.

FTRL FOR OLO: THEORETICAL GUARANTEES

- Interpretation of the terms in the proposition, i.e., of

$$R_T^{FTRL}(\tilde{a}) \leq \frac{1}{2\eta}||\tilde{a}||_2^2 + \eta\sum_{t=1}^{T}||z_t||_2^2 :$$

  - $||\tilde{a}||_2^2$ represents a *bias term:* The regret upper bound of FTRL is always biased by the term $||\tilde{a}||_2^2$. The impact of the bias term can be reduced by a higher regularization magnitude, i.e., a higher choice of η.
  - $\sum_{t=1}^{T}||z_t||_2^2$ represents a *"variance" term*: The more the environment data z_t varies, the larger this term. Hence, for a high variance a smaller regularization magnitude is needed, i.e., a smaller choice of η.

- Thus, we have a trade-off for the optimal choice of η : Making η large, leads to a smaller bias but at the expense of a higher variance and making η small leads to a smaller variance at the expense of a higher bias.

$\Rightarrow$ With the right choice of η, we can prevent the instability of FTRL for an online linear optimization (OLO) problem.

FTRL FOR OLO: THEORETICAL GUARANTEES

- Under certain assumptions we can balance the trade-off induced by the bias and the variance by choosing η appropriately.

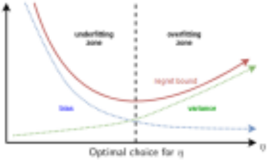
Interpretation of the terms in the proposition, i.e., of $R^{FTRL}(\tilde{a}) \leq \frac{1}{2\eta}||\tilde{a}||_2^2 + \eta$

  - $||\tilde{a}||_2^2$ represents a *bias term:* The regret upper bound of FTRL is always biased by the term $||\tilde{a}||_2^2$. The impact of the bias term can be reduced by a higher regularization magnitude, i.e., a higher choice of η.
  - represents a : The more the environment data z_t varies, the larger this term. Hence, for a high variance a smaller regularization magnitude is needed, i.e., a smaller choice of η.
- Thus, we have a trade-off for the optimal choice of η : Making η large, leads to a smaller bias but at the expense of a higher and making η small leads to a smaller at the expense of a higher bias.

$\Rightarrow$ With the right choice of η, we can prevent the instability of FTRL for an online linear optimization (OLO) problem.

FTRL FOR OLO: THEORETICAL GUARANTEES

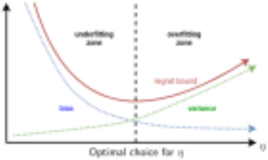
- Under certain assumptions we can balance the trade-off induced by the bias and the variance by choosing η appropriately.



- **Corollary:** Suppose we use the FTRL algorithm with the squared L2-norm regularization on an online linear optimization problem with $\mathcal{A} \subset \mathbb{R}^d$ such that
  - $\sup_{\tilde{a} \in \mathcal{A}} ||\tilde{a}||_2 \leq B$ for some finite constant $B > 0$,
  - $\sup_{z \in \mathcal{Z}} ||z||_2 \leq V$ for some finite constant $V > 0$.

FTRL FOR OLO: THEORETICAL GUARANTEES

- Under certain assumptions we can balance the trade-off induced by the bias and the variance by choosing η appropriately.



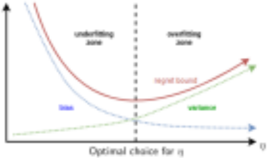
- **Corollary:** Suppose we use the FTRL algorithm with the squared L2-norm regularization on an online linear optimization problem with $\mathcal{A} \subset \mathbb{R}^d$ such that
  - $\sup_{\tilde{a} \in \mathcal{A}} ||\tilde{a}||_2 \leq B$ for some finite constant $B > 0$,
  - $\sup_{z \in \mathcal{Z}} ||z||_2 \leq V$ for some finite constant $V > 0$.

  Then, by choosing the step size η for FTRL as $\eta = \frac{B}{V\sqrt{2T}}$ it holds that

$$R_T^{FTRL} \leq BV\sqrt{2T}.$$

FTRL FOR OLO: THEORETICAL GUARANTEES

- Under certain assumptions we can balance the trade-off induced by the bias and the variance by choosing η appropriately.



- **Corollary:** Suppose we use the FTRL algorithm with the squared L2-norm regularization on an online linear optimization problem with $\mathcal{A} \subset \mathbb{R}^d$ such that
  - $\sup_{\tilde{a} \in \mathcal{A}} ||\tilde{a}||_2 \leq B$ for some finite constant $B > 0$,
  - $\sup_{z \in \mathcal{Z}} ||z||_2 \leq V$ for some finite constant $V > 0$.

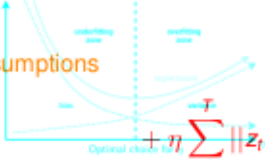
  Then, by choosing the step size η for FTRL as $\eta = \frac{B}{V\sqrt{2T}}$ it holds that

$$R_T^{FTRL} \leq BV\sqrt{2T}.$$

- Note that the (optimal) parameter η depends on the time horizon T, which is oftentimes not known in advance. However, there are some tricks (i.e., the *doubling trick*), which can help in such cases.

FTRL FOR OLO: THEORETICAL GUARANTEES

- Under certain assumptions we can balance the trade-off induced by the bias and the variance by choosing η appropriately.



- **Proof:**
  - By the latter proposition and the assumptions

$$R_T^{FTRL}(\tilde{a}) \leq \frac{1}{2\eta}||\tilde{a}||_2^2 + \eta \sum_{t=1}^{T} ||z_t||_2^2$$

$$\leq \frac{B^2}{2\eta} + \eta T V^2.$$

- **Corollary:** Suppose we use the FTRL algorithm with the squared L2-norm regularization on an online linear optimization problem with $\mathcal{A} \subset \mathbb{R}^d$ such that
  - $\sup_{\tilde{a} \in \mathcal{A}} ||\tilde{a}||_2 \leq B$ for some finite constant $B > 0$,
  - $\sup_{z \in \mathcal{Z}} ||z||_2 \leq V$ for some finite constant $V > 0$.

  Then, by choosing the step size η for FTRL as $\eta = \frac{B}{V\sqrt{2T}}$ it holds that

$$R_T^{FTRL} \leq BV\sqrt{2T}.$$

- Note that the (optimal) parameter η depends on the time horizon T, which is oftentimes not known in advance. However, there are some tricks (i.e., the *doubling trick*), which can help in such cases.

FTRL FOR OLO: THEORETICAL GUARANTEES

- **Proof:**
  - By the latter proposition and the assumptions

$$
\begin{aligned}
R_T^{FTRL}(\tilde{a}) &\leq \frac{1}{2\eta}\,||\tilde{a}||_2^2 && + \eta \sum_{t=1}^{T} ||z_t||_2^2 \\
&\leq \frac{B^2}{2\eta} && + \eta\, T\, V^2 .
\end{aligned}
$$

  - The right-hand side of the latter display is independent of $\tilde{a}$, so that

$$R_T^{FTRL} \leq \frac{B^2}{2\eta} + \eta\, T\, V^2 .$$

FTRL FOR OLO: THEORETICAL GUARANTEES

- **Proof:**
  - By the latter proposition and the assumptions

$$
\begin{aligned}
R_T^{FTRL}(\tilde{a}) &\leq \frac{1}{2\eta}\,||\tilde{a}||_2^2 && + \eta \sum_{t=1}^{T} ||z_t||_2^2 \\
&\leq \frac{B^2}{2\eta} && + \eta\, T\, V^2 .
\end{aligned}
$$

  - The right-hand side of the latter display is independent of $\tilde{a}$, so that

$$
R_T^{FTRL} \leq \frac{B^2}{2\eta} + \eta\, T\, V^2 .
$$

  - Now, the right-hand side of the latter display is a function of the form $f(\eta) = a/\eta + b\eta$ for some suitable $a, b > 0$.

FTRL FOR OLO: THEORETICAL GUARANTEES

- **Proof:**
  - By the latter proposition and the assumptions

$$R_T^{FTRL}(\tilde{a}) \leq \frac{1}{2\eta}||\tilde{a}||_2^2 + \eta\sum_{t=1}^{T}||z_t||_2^2 \leq \frac{B^2}{2\eta} + \eta\, T\, V^2.$$

  - The right-hand side of the latter display is independent of $\tilde{a}$, so that

$$R_T^{FTRL} \leq \frac{B^2}{2\eta} + \eta\, T\, V^2.$$

  - Now, the right-hand side of the latter display is a function of the form $f(\eta) = a/\eta + b\eta$ for some suitable $a, b > 0$.
  - Minimizing f with respect to η results in the minimizer $\eta^* = \frac{B}{V\sqrt{2T}}$.

FTRL FOR OLO: THEORETICAL GUARANTEES

- **Proof:**
  - By the latter proposition and the assumptions

$$
\begin{aligned}
R_T^{FTRL}(\tilde{a}) &\leq \frac{1}{2\eta}\,||\tilde{a}||_2^2 && + \eta \sum_{t=1}^{T} ||z_t||_2^2 \\
&\leq \frac{B^2}{2\eta} && + \eta\, T\, V^2 .
\end{aligned}
$$

  - The right-hand side of the latter display is independent of $\tilde{a}$, so that

$$R_T^{FTRL} \leq \frac{B^2}{2\eta} + \eta\, T\, V^2 .$$

  - Now, the right-hand side of the latter display is a function of the form $f(\eta) = a/\eta + b\eta$ for some suitable $a, b > 0$.
  - Minimizing f with respect to η results in the minimizer $\eta^* = \frac{B}{V\sqrt{2T}}$.
  - Plugging this minimizer into the latter display leads to the asserted inequality. □

DESIRED RESULTS

- With the FTRL algorithm we can cope with
  - online quadratic optimization (OQO) problems by using no regularity ($\psi \equiv 0$). In this case, we have satisfactory regret guarantees and also a quick update rule for a_{t+1}^{FTRL} (It is just the empirical average over all data points seen till t),

THEORETICAL GUARANTEES

- Proof:
  - By the latter proposition and the assumptions

$$R_T^{FTRL}(\tilde{a}) \le \frac{1}{2\eta}\|\tilde{a}\|_2^2 + \eta \sum_{t=1}^{T} \|z_t\|_2^2 \le \frac{B^2}{2\eta} + \eta\, T\, V^2.$$

  - The right-hand side of the latter display is independent of $\tilde{a}$, so that

$$R_T^{FTRL} \le \frac{B^2}{2\eta} + \eta\, T\, V^2.$$

  - Now, the right-hand side of the latter display is a function of the form $f(\eta) = a/\eta + b\eta$ for some suitable $a, b > 0$.
  - Minimizing f with respect to η results in the minimizer $\eta^* = \frac{B}{V\sqrt{2T}}$.
  - Plugging this minimizer into the latter display leads to the asserted inequality. □

DESIRED RESULTS

- With the FTRL algorithm we can cope with
  - online quadratic optimization (OQO) problems by using no regularity ($\psi \equiv 0$). In this case, we have satisfactory regret guarantees and also a quick update rule for a_{t+1}^{FTRL} (It is just the empirical average over all data points seen till t),
  - online linear optimization (OLO) problems by using a suitable regularization function. In this case, we have quick update formulas and satisfactory regret guarantees as well.

DESIRED RESULTS

- With the FTRL algorithm we can cope with
  - online quadratic optimization (OQO) problems by using no regularity ($\psi \equiv 0$). In this case, we have satisfactory regret guarantees and also a quick update rule for a_{t+1}^{FTRL} (It is just the empirical average over all data points seen till t),
  - online linear optimization (OLO) problems by using a suitable regularization function. In this case, we have quick update formulas and satisfactory regret guarantees as well.

$\Rightarrow$ But what about other online learning problems or rather other loss functions?

- What we wish to have is an approach such that we can achieve for a large class of loss functions the advantages of FTRL for OLO and OCO problems:

  (a) reasonable regret upper bounds;
  (b) a quick update formula.

DESIRED RESULTS

- With the FTRL algorithm we can cope with
  - online quadratic optimization (OQO) problems by using no regularity ($\psi \equiv 0$). In this case, we have satisfactory regret guarantees and also a quick update rule for a_{t+1}^{FTRL} (It is just the empirical average over all data points seen till t),
  - online linear optimization (OLO) problems by using a suitable regularization function. In this case, we have quick update formulas and satisfactory regret guarantees as well.

$\Rightarrow$ But what about other online learning problems or rather other loss functions?

- What we wish to have is an approach such that we can achieve for a large class of loss functions L the advantages of FTRL for OLO and OCO problems:

  (a) reasonable regret upper bounds;
  (b) a quick update formula.