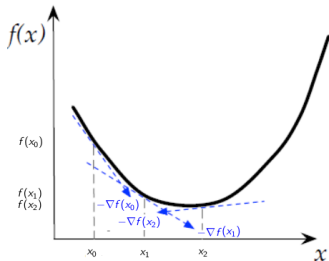


# Advanced Machine Learning

## Online Convex Optimization - Part 2 -



### Learning goals

- Know the connection between OGD and FTRL via linearization of convex functions
- See how this implies regret bounds for OGD
- Get to know the theoretical limits for online convex optimization

# ONLINE GRADIENT DESCENT

- The *Online Gradient Descent* (OGD) algorithm with step size  $\eta > 0$  chooses its action by

$$a_{t+1}^{\text{OGD}} = a_t^{\text{OGD}} - \eta \nabla_a L(a_t^{\text{OGD}}, z_t), \quad t = 1, \dots, T. \quad (1)$$

(Technical side note: For this update formula we assume that  $\mathcal{A} = \mathbb{R}^d$ . Moreover, the first action  $a_1^{\text{OGD}}$  is arbitrary. )



# ONLINE GRADIENT DESCENT

- The *Online Gradient Descent* (OGD) algorithm with step size  $\eta > 0$  chooses its action by

$$a_{t+1}^{\text{OGD}} = a_t^{\text{OGD}} - \eta \nabla_a L(a_t^{\text{OGD}}, z_t), \quad t = 1, \dots, T. \quad (1)$$

(Technical side note: For this update formula we assume that  $\mathcal{A} = \mathbb{R}^d$ . Moreover, the first action  $a_1^{\text{OGD}}$  is arbitrary. )

- We have the following connection between FTRL and OGD:
  - Let  $\tilde{z}_t^{\text{OGD}} := \nabla_a L(a_t^{\text{OGD}}, z_t)$  for any  $t = 1, \dots, T$ .
  - The update formula for FTRL with  $L_2$  norm regularization for the linear loss  $L^{\text{lin}}$  and the environmental data  $\tilde{z}_t^{\text{OGD}}$  is

$$a_{t+1}^{\text{FTRL}} = a_t^{\text{FTRL}} - \eta \tilde{z}_t^{\text{OGD}} = a_t^{\text{FTRL}} - \eta \nabla_a L(a_t^{\text{OGD}}, z_t).$$

- If we have that  $a_1^{\text{FTRL}} = a_1^{\text{OGD}}$ , then it iteratively follows that  $a_{t+1}^{\text{FTRL}} = a_{t+1}^{\text{OGD}}$  for any  $t = 1, \dots, T$  in this case.



# ONLINE GRADIENT DESCENT: DEFINITION AND PROPERTIES

- With the deliberations above we can infer that

$$\begin{aligned} R_{T,L}^{\text{OGD}}(\tilde{a} \mid (z_t)_t) &= \sum_{t=1}^T L(a_t^{\text{OGD}}, z_t) - L(\tilde{a}, z_t) \\ &\leq \sum_{t=1}^T L^{\text{lin}}(a_t^{\text{OGD}}, \tilde{z}_t^{\text{OGD}}) - L^{\text{lin}}(\tilde{a}, \tilde{z}_t^{\text{OGD}}) \\ &\quad (\text{if } a_1^{\text{OGD}} = a_1^{\text{FTRL}}) \sum_{t=1}^T L^{\text{lin}}(a_t^{\text{FTRL}}, \tilde{z}_t^{\text{OGD}}) - L^{\text{lin}}(\tilde{a}, \tilde{z}_t^{\text{OGD}}) \\ &= R_{T,L^{\text{lin}}}^{\text{FTRL}}(\tilde{a} \mid (\tilde{z}_t^{\text{OGD}})_t), \end{aligned}$$

where we write in the subscripts of the regret the corresponding loss function and also include the corresponding environmental data as a second argument in order to emphasize the connections.



# ONLINE GRADIENT DESCENT: DEFINITION AND PROPERTIES

- With the deliberations above we can infer that

$$\begin{aligned} R_{T,L}^{\text{OGD}}(\tilde{a} \mid (z_t)_t) &= \sum_{t=1}^T L(a_t^{\text{OGD}}, z_t) - L(\tilde{a}, z_t) \\ &\leq \sum_{t=1}^T L^{\text{lin}}(a_t^{\text{OGD}}, \tilde{z}_t^{\text{OGD}}) - L^{\text{lin}}(\tilde{a}, \tilde{z}_t^{\text{OGD}}) \\ &\quad (\text{if } a_1^{\text{OGD}} = a_1^{\text{FTRL}}) \sum_{t=1}^T L^{\text{lin}}(a_t^{\text{FTRL}}, \tilde{z}_t^{\text{OGD}}) - L^{\text{lin}}(\tilde{a}, \tilde{z}_t^{\text{OGD}}) \\ &= R_{T,L^{\text{lin}}}^{\text{FTRL}}(\tilde{a} \mid (\tilde{z}_t^{\text{OGD}})_t), \end{aligned}$$

where we write in the subscripts of the regret the corresponding loss function and also include the corresponding environmental data as a second argument in order to emphasize the connections.

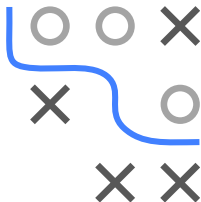
- *Interpretation:* The regret of the FTRL algorithm (with  $L_2$  norm regularization) for the online linear optimization problem (characterized by the linear loss  $L^{\text{lin}}$ ) with environmental data  $\tilde{z}_t^{\text{OGD}}$  is an upper bound for the OGD algorithm for the online convex problem (characterized by a differentiable convex loss  $L$ ) with the original environmental data  $z_t$ .



# ONLINE GRADIENT DESCENT: REGRET

- Due to this connection we immediately obtain a similar decomposition of the regret upper bound into a bias term and a variance term as for the FTRL algorithm for OLO problems.
- **Corollary.** Using the OGD algorithm on any online convex optimization problem (with differentiable loss function  $L$ ) leads to a regret of OGD with respect to any action  $\tilde{a} \in \mathcal{A}$  of

$$\begin{aligned} R_T^{\text{OGD}}(\tilde{a}) &\leq \frac{1}{2\eta} \|\tilde{a}\|_2^2 + \eta \sum_{t=1}^T \|\tilde{z}_t^{\text{OGD}}\|_2^2 \\ &= \frac{1}{2\eta} \|\tilde{a}\|_2^2 + \eta \sum_{t=1}^T \|\nabla_a L(a_t^{\text{OGD}}, z_t)\|_2^2. \end{aligned}$$



# ONLINE GRADIENT DESCENT: REGRET

- Due to this connection we immediately obtain a similar decomposition of the regret upper bound into a bias term and a variance term as for the FTRL algorithm for OLO problems.
- **Corollary.** Using the OGD algorithm on any online convex optimization problem (with differentiable loss function  $L$ ) leads to a regret of OGD with respect to any action  $\tilde{a} \in \mathcal{A}$  of

$$\begin{aligned} R_T^{\text{OGD}}(\tilde{a}) &\leq \frac{1}{2\eta} \|\tilde{a}\|_2^2 + \eta \sum_{t=1}^T \|\tilde{z}_t^{\text{OGD}}\|_2^2 \\ &= \frac{1}{2\eta} \|\tilde{a}\|_2^2 + \eta \sum_{t=1}^T \|\nabla_a L(a_t^{\text{OGD}}, z_t)\|_2^2. \end{aligned}$$

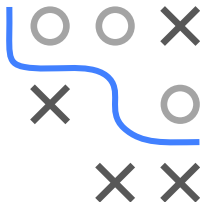
- Note that the step size  $\eta > 0$  of OGD has the same role as the regularization magnitude of FTRL: It should balance the trade-off between the bias- and the variance-term.



# ONLINE GRADIENT DESCENT: REGRET

- As a consequence, we can also derive a similar order of the regret for the OGD algorithm on OCO problems as for the FTRL on OLO problems by imposing a slightly different assumption on the (new) “variance” term

$$\sum_{t=1}^T \|\nabla_a L(a_t^{\text{OGD}}, z_t)\|_2^2.$$





# ONLINE GRADIENT DESCENT: REGRET

- As a consequence, we can also derive a similar order of the regret for the OGD algorithm on OCO problems as for the FTRL on OLO problems by imposing a slightly different assumption on the (new) “variance” term

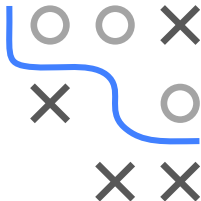
$$\sum_{t=1}^T \|\nabla_a L(a_t^{\text{OGD}}, z_t)\|_2^2.$$

- Corollary:** Suppose we use the OGD algorithm on an online convex optimization problem with a convex action space  $\mathcal{A} \subset \mathbb{R}^d$  such that

- $\sup_{\tilde{a} \in \mathcal{A}} \|\tilde{a}\|_2 \leq B$  for some finite constant  $B > 0$
- $\sup_{a \in \mathcal{A}, z \in \mathcal{Z}} \|\nabla_a L(a, z)\|_2 \leq V$  for some finite constant  $V > 0$ .

Then, by choosing the step size  $\eta$  for OGD as  $\eta = \frac{B}{V\sqrt{2T}}$  we get

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



# REGRET LOWER BOUNDS FOR OCO

- **Theorem.** For any online learning algorithm there exists an online convex optimization problem characterized by
  - a convex loss function  $L$ ,



# REGRET LOWER BOUNDS FOR OCO

- **Theorem.** For any online learning algorithm there exists an online convex optimization problem characterized by
  - a convex loss function  $L$ ,
  - a bounded (convex) action space  $\mathcal{A} = [-B, B]^d$  for some finite constant  $B > 0$ ,



# REGRET LOWER BOUNDS FOR OCO

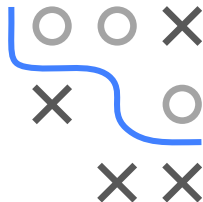
- **Theorem.** For any online learning algorithm there exists an online convex optimization problem characterized by
  - a convex loss function  $L$ ,
  - a bounded (convex) action space  $\mathcal{A} = [-B, B]^d$  for some finite constant  $B > 0$ ,
  - and bounded gradients  $\sup_{a \in \mathcal{A}, z \in \mathcal{Z}} \|\nabla_a L(a, z)\|_2 \leq V$  for some finite constant  $V > 0$ ,



# REGRET LOWER BOUNDS FOR OCO

- **Theorem.** For any online learning algorithm there exists an online convex optimization problem characterized by
  - a convex loss function  $L$ ,
  - a bounded (convex) action space  $\mathcal{A} = [-B, B]^d$  for some finite constant  $B > 0$ ,
  - and bounded gradients  $\sup_{a \in \mathcal{A}, z \in \mathcal{Z}} \|\nabla_a L(a, z)\|_2 \leq V$  for some finite constant  $V > 0$ ,

such that the algorithm incurs a regret of  $\Omega(\sqrt{T})$  in the worst case.

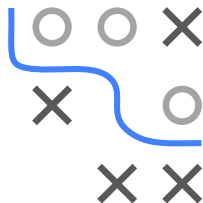


# REGRET LOWER BOUNDS FOR OCO

- **Theorem.** For any online learning algorithm there exists an online convex optimization problem characterized by
  - a convex loss function  $L$ ,
  - a bounded (convex) action space  $\mathcal{A} = [-B, B]^d$  for some finite constant  $B > 0$ ,
  - and bounded gradients  $\sup_{a \in \mathcal{A}, z \in \mathcal{Z}} \|\nabla_a L(a, z)\|_2 \leq V$  for some finite constant  $V > 0$ ,

such that the algorithm incurs a regret of  $\Omega(\sqrt{T})$  in the worst case.

- Recall that under (almost) the same assumptions as the theorem above, we have  $R_T^{\text{OGD}} \leq BV\sqrt{2T}$ .



# REGRET LOWER BOUNDS FOR OCO

- **Theorem.** For any online learning algorithm there exists an online convex optimization problem characterized by
  - a convex loss function  $L$ ,
  - a bounded (convex) action space  $\mathcal{A} = [-B, B]^d$  for some finite constant  $B > 0$ ,
  - and bounded gradients  $\sup_{a \in \mathcal{A}, z \in \mathcal{Z}} \|\nabla_a L(a, z)\|_2 \leq V$  for some finite constant  $V > 0$ ,

such that the algorithm incurs a regret of  $\Omega(\sqrt{T})$  in the worst case.

- Recall that under (almost) the same assumptions as the theorem above, we have  $R_T^{\text{OGD}} \leq BV\sqrt{2T}$ .

↪ This result shows that the Online Gradient Descent is *optimal* regarding its order of its regret with respect to the time horizon  $T$ .

