ONLINE GRADIENT DESCENT

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

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

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

ONLINE GRADIENT DESCENT

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

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

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

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

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

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

ONLINE GRADIENT DESCENT: DEFINITION AND PROPERTIES

- With the deliberations above we can infer that

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

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

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

  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^{lin}) with environmental data $\tilde{z}_t^{\mathrm{OGD}}$ 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) leads to a regret of OGD with respect to any action $\tilde{a} \in \mathcal{A}$ of

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

$$= \frac{1}{2\eta} \|\tilde{a}\|_2^2 + \eta \sum_{t=1}^{T} \left\| \nabla_a \ell(a_t^{\mathrm{OGD}}, z_t) \right\|_2^2 .$$

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) leads to a regret of OGD with respect to any action $\tilde{a} \in \mathcal{A}$ of

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

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

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

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

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

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

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 $\sup_{a \in \mathcal{A}, z \in \mathcal{Z}} ||\nabla_a \ell(a, z)||_2 \leq V$ 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 $R_T^{\mathrm{OGD}} \leq BV\sqrt{2T}$.

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

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