ONLINE GRADIENT DESCENT

 The Online Gradient Descent (OGD) algorithm with step size η > 0 chooses its action by

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

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

× < 0 × × ×

ONLINE GRADIENT DESCENT

 The Online Gradient Descent (OGD) algorithm with step size η > 0 chooses its action by

$$a_{t+1}^{\text{OGD}} = a_t^{\text{OGD}} - \eta \nabla_a(a_t^{\text{OGD}}, z_t), \quad t = 1, \dots, T.$$
 (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 $\tilde{z}_t^{\text{DGD}} := \nabla_a(a_t^{\text{DGD}}, z_t)$ for any $t = 1, \dots, T$.
 - The update formula for FTRL with 2 norm regularization for the linear loss L^{lin} and the environmental data ž^{0GD}_t is

$$a_{t+1}^{ ext{FTRL}} = a_t^{ ext{FTRL}} - \eta \widetilde{z}_t^{ ext{OGD}} = a_t^{ ext{FTRL}} - \eta
abla_a(a_t^{ ext{OGD}}, z_t).$$

• 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

$$egin{aligned} &\mathcal{R}_{T,}^{ ext{OGD}}(ilde{a} \mid (z_t)_t) = \sum_{t=1}^T (a_t^{ ext{OGD}}, z_t) - (ilde{a}, z_t) \ &\leq \sum_{t=1}^T \mathcal{L}^{ ext{lin}}(a_t^{ ext{OGD}}, ilde{z}_t^{ ext{OGD}}) - \mathcal{L}^{ ext{lin}}(ilde{a}, ilde{z}_t^{ ext{OGD}}) \ &(ext{if } a_1^{ ext{OGD}} = a_1^{ ext{FTRL}}) \sum_{t=1}^T \mathcal{L}^{ ext{lin}}(a_t^{ ext{FTRL}}, ilde{z}_t^{ ext{OGD}}) - \mathcal{L}^{ ext{lin}}(ilde{a}, ilde{z}_t^{ ext{OGD}}) \ &= \mathcal{R}_{T, \mathcal{L}^{ ext{lin}}}^{ ext{FTRL}}(ilde{a} \mid (ilde{z}_t^{ ext{OGD}})_t), \end{aligned}$$

× × 0 × × ×

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

$$\begin{split} \mathcal{R}_{T,}^{\text{OGD}}(\tilde{a} \mid (z_{t})_{t}) &= \sum_{t=1}^{T} (a_{t}^{\text{OGD}}, z_{t}) - (\tilde{a}, z_{t}) \\ &\leq \sum_{t=1}^{T} \mathcal{L}^{\text{lin}}(a_{t}^{\text{OGD}}, \tilde{z}_{t}^{\text{OGD}}) - \mathcal{L}^{\text{lin}}(\tilde{a}, \tilde{z}_{t}^{\text{OGD}}) \\ &(\text{if } a_{1}^{\text{OGD}} = a_{1}^{\text{FTRL}}) \sum_{t=1}^{T} \mathcal{L}^{\text{lin}}(a_{t}^{\text{FTRL}}, \tilde{z}_{t}^{\text{OGD}}) - \mathcal{L}^{\text{lin}}(\tilde{a}, \tilde{z}_{t}^{\text{OGD}}) \\ &= \mathcal{R}_{T, \mathcal{L}^{\text{lin}}}^{\text{FTRL}}(\tilde{a} \mid (\tilde{z}_{t}^{\text{OGD}})_{t}), \end{split}$$

× × 0 × × ×

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 $_2$ norm regularization) for the online linear optimization problem (characterized by the linear loss L^{1in}) with environmental data \tilde{z}_t^{00D} is an upper bound for the OGD algorithm for the online convex problem (characterized by a differentiable convex loss) with the original environmental data z_t .

- 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 $\tilde{a} \in A$ of

$$\begin{split} \mathcal{R}_{T}^{\texttt{DGD}}(\tilde{a}) &\leq \frac{1}{2\eta} \left| |\tilde{a}||_{2}^{2} + \eta \sum_{t=1}^{T} \left| \left| \tilde{z}_{t}^{\texttt{DGD}} \right| \right|_{2}^{2} \\ &= \frac{1}{2\eta} \left| |\tilde{a}||_{2}^{2} + \eta \sum_{t=1}^{T} \left| \left| \nabla_{a}(a_{t}^{\texttt{DGD}}, z_{t}) \right| \right|_{2}^{2} \end{split}$$

- 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 $\tilde{a} \in A$ of

$$\begin{split} \mathcal{R}_{T}^{\texttt{OGD}}(\tilde{a}) &\leq \frac{1}{2\eta} \left| |\tilde{a}||_{2}^{2} + \eta \sum_{t=1}^{T} \left| \left| \tilde{z}_{t}^{\texttt{OGD}} \right| \right|_{2}^{2} \\ &= \frac{1}{2\eta} \left| |\tilde{a}||_{2}^{2} + \eta \sum_{t=1}^{T} \left| \left| \nabla_{a}(a_{t}^{\texttt{OGD}}, z_{t}) \right| \right|_{2}^{2} \end{split}$$

• 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.

× × 0 × × ×

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

- 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 $\sum_{t=1}^{T} ||\nabla_a(a_t^{\text{OGD}}, z_t)||_2^2$.
- **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
 - $\sup_{a \in \mathcal{A}, z \in \mathcal{Z}} ||\nabla_a(a, z)||_2 \leq V$ for some finite constant V > 0.

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

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

× 0 0 × × ×

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

× 0 0 × 0 × ×

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

× 0 0 × 0 × ×

- **Theorem.** For any online learning algorithm there exists an online convex optimization problem characterized by
 - a convex loss function ,
 - a bounded (convex) action space A = [-B, B]^d for some finite constant B > 0,
 - and bounded gradients sup_{a∈A,z∈Z} ||∇_a(a, z)||₂ ≤ V for some finite constant V > 0,

× × 0 × × ×

- **Theorem.** For any online learning algorithm there exists an online convex optimization problem characterized by
 - a convex loss function ,
 - a bounded (convex) action space A = [−B, B]^d for some finite constant B > 0,
 - and bounded gradients sup_{a∈A,z∈Z} ||∇_a(a, z)||₂ ≤ V for some finite constant V > 0,

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

× 0 0 × × ×

- **Theorem.** For any online learning algorithm there exists an online convex optimization problem characterized by
 - a convex loss function ,
 - a bounded (convex) action space A = [-B, B]^d for some finite constant B > 0,
 - and bounded gradients sup_{a∈A,z∈Z} ||∇_a(a, z)||₂ ≤ 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 $R_T^{\text{DGD}} \leq BV\sqrt{2T}$.

- **Theorem.** For any online learning algorithm there exists an online convex optimization problem characterized by
 - a convex loss function ,
 - a bounded (convex) action space A = [-B, B]^d for some finite constant B > 0,
 - and bounded gradients sup_{a∈A,z∈Z} ||∇_a(a, z)||₂ ≤ 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 $R_T^{\text{DGD}} \leq BV\sqrt{2T}$.
- \rightsquigarrow This result shows that the Online Gradient Descent is *optimal* regarding its order of its regret with respect to the time horizon *T*.

× 0 0 × × ×