Online Convex Optimization

An overview of algorithms and techniques

Fan Yao, University of Virginia

Outline

- Problem Set-up
- Follow the Regularized Leader (FTRL)
- Online Projected Sub-gradient Descent (PSGD)
- Exponentiated gradient (EG)
- EXP-3 and it's variants
- Online Mirror Descent (OMD)
- Dual Averaging (DA)

Set-Up

Online Convex Optimization Problem

- Convex set C.
- For $t = 1$ to T do
  - predict $\mathbf{w}_t \in C$.
  - receive convex loss function $f_t \colon C \to \mathbb{R}$.
  - incur loss $f_t(\mathbf{w}_t)$.
- Regret of algorithm $\mathcal{A}$:

$$R_T(\mathcal{A}) = \sum_{t=1}^{T} f_t(\mathbf{w}_t) - \inf_{\mathbf{w} \in C} \sum_{t=1}^{T} f_t(\mathbf{w}).$$

Set-Up

Online Convex Optimization Problem

- Convex set C.
- For $t = 1$ to T do
  - predict $\mathbf{w}_t \in C$.
  - receive convex loss function $f_t : C \to \mathbb{R}$.
  - incur loss $f_t(\mathbf{w}_t)$.
- Regret of algorithm $\mathcal{A}$:

$$R_T(\mathcal{A}) = \sum_{t=1}^{T} f_t(\mathbf{w}_t) - \inf_{\mathbf{w} \in C} \sum_{t=1}^{T} f_t(\mathbf{w}).$$

Remarks:

1. In addition to convexity, Lipschitz continuity is often assumed for f_t .
2. It is standard convex optimization if all f_t takes the same form.
3. It is also common to use l_t to refer to the loss function. We will use the notations $l_t(x_t) \leftrightarrow f_t(w_t)$ interchangeably.

Feedback Assumptions

- Full information — Given f_t, optimizer can evaluate $f_t(w), \forall w \in C$
- Bandit information
  - First order feedback — Given f_t, optimizer can evaluate $f_t(w_t), \nabla f_t(w_t)$
  - Zeroth order feedback — Given f_t, optimizer can only evaluate $f_t(w_t)$

Feedback Assumptions

- Full information Given f_t, optimizer can evaluate $f_t(w), \forall w \in C$
- Bandit information
  - First order feedback Given f_t, optimizer can evaluate $f_t(w_t), \nabla f_t(w_t)$
  - Zeroth order feedback Given f_t, optimizer can only evaluate $f_t(w_t)$

Under different feedback assumptions, we are interested in developing no-regret learning algorithms, i.e.,

$$R_T(\mathscr{A}) = o(T).$$

Full Information Feedback

Follow the Leader (FTL)

- In the full information feedback setting, FTL could be a plausible choice:

$$x_{t+1} \in \arg\min_{x \in C} \sum_{s=1}^{t} l_s(x) . \qquad (FTL\ update\ rule)$$

Follow the Leader (FTL)

- In the full information feedback setting, FTL could be a plausible choice:

$$x_{t+1} \in \arg\min_{x \in C} \sum_{s=1}^{t} l_s(x) \, . \qquad \textit{(FTL update rule)}$$

- However, FTL has O(T) regret even for linear loss functions:

$$\ell_t(x) = \begin{cases} -x/2 & \text{for } t = 1, \\ x & \text{if } t > 1 \text{ is even}, \\ -x & \text{if } t > 1 \text{ is odd}. \end{cases}$$

Follow the Leader (FTL)

- In the full information feedback setting, FTL could be a plausible choice:

$$x_{t+1} \in \arg\min_{x \in C} \sum_{s=1}^{t} l_s(x) . \qquad (FTL\ update\ rule)$$

- However, FTL has O(T) regret even for linear loss functions:

$$\ell_t(x) = \begin{cases} -x/2 & \text{for } t = 1, \\ x & \text{if } t > 1 \text{ is even,} \\ -x & \text{if } t > 1 \text{ is odd.} \end{cases}$$

- FTL is too aggressive. Need to impose restrictions on $\{x_t\}$ to avoid jiggling.

Follow the Regularized Leader (FTRL)

- Solution: introducing a regularization term $h(x)$:

$$x_{t+1} \in \arg\min_{x \in C} \left\{ \sum_{s=1}^{t} l_s(x) + \frac{1}{\gamma} h(x) \right\}. \qquad (\textit{FTRL update rule})$$

Follow the Regularized Leader (FTRL)

- Solution: introducing a regularization term $h(x)$:

$$x_{t+1} \in \arg\min_{x \in C} \left\{ \sum_{s=1}^{t} l_s(x) + \frac{1}{\gamma} h(x) \right\}. \qquad (FTRL\ update\ rule)$$

- $h(x)$ is continuous and strongly convex, i.e., $\exists K > 0, s.t.$

$$[\lambda h(x') + (1-\lambda) h(x)] - h(\lambda x' + (1-\lambda) x) \geq \frac{K}{2} \lambda (1-\lambda) \| x' - x \|^2$$

for all $\lambda \in [0,1], x, x' \in C$.

Regret of FTRL [Shalev-Shwartz, 2007]

- If $h(x)$ is continuous and strongly convex, and each l_t is convex and Lipschitz continuous with universal Lipschitz constant L, $H = \max_{x \in C} h(x) - \min_{x \in C} h(x)$ is the depth of h over C. Then, the regret of FTRL can be bounded by

$$R_T(\text{FTRL}) \leq 2L\sqrt{(H/K)T}\,.$$

Regret of FTRL [Shalev-Shwartz, 2007]

- If $h(x)$ is continuous and strongly convex, and each l_t is convex and Lipschitz continuous with universal Lipschitz constant L, $H = \max_{x \in C} h(x) - \min_{x \in C} h(x)$ is the depth of h over C. Then, the regret of FTRL can be bounded by

$$R_T(\text{FTRL}) \leq 2L\sqrt{(H/K)T}\,.$$

Remarks:

1. FTL and FTRL are closely related to the learning policies known in economics and game theory as fictitious play (FP) and smooth fictitious play (SFP), respectively. These policies correspond to playing a best response (resp. regularized or smooth best response) to the empirical history of play of one's opponents.
2. Our assumptions: the optimizer has full information access to the loss functions, and the minimization sub-problem can be solved efficiently.

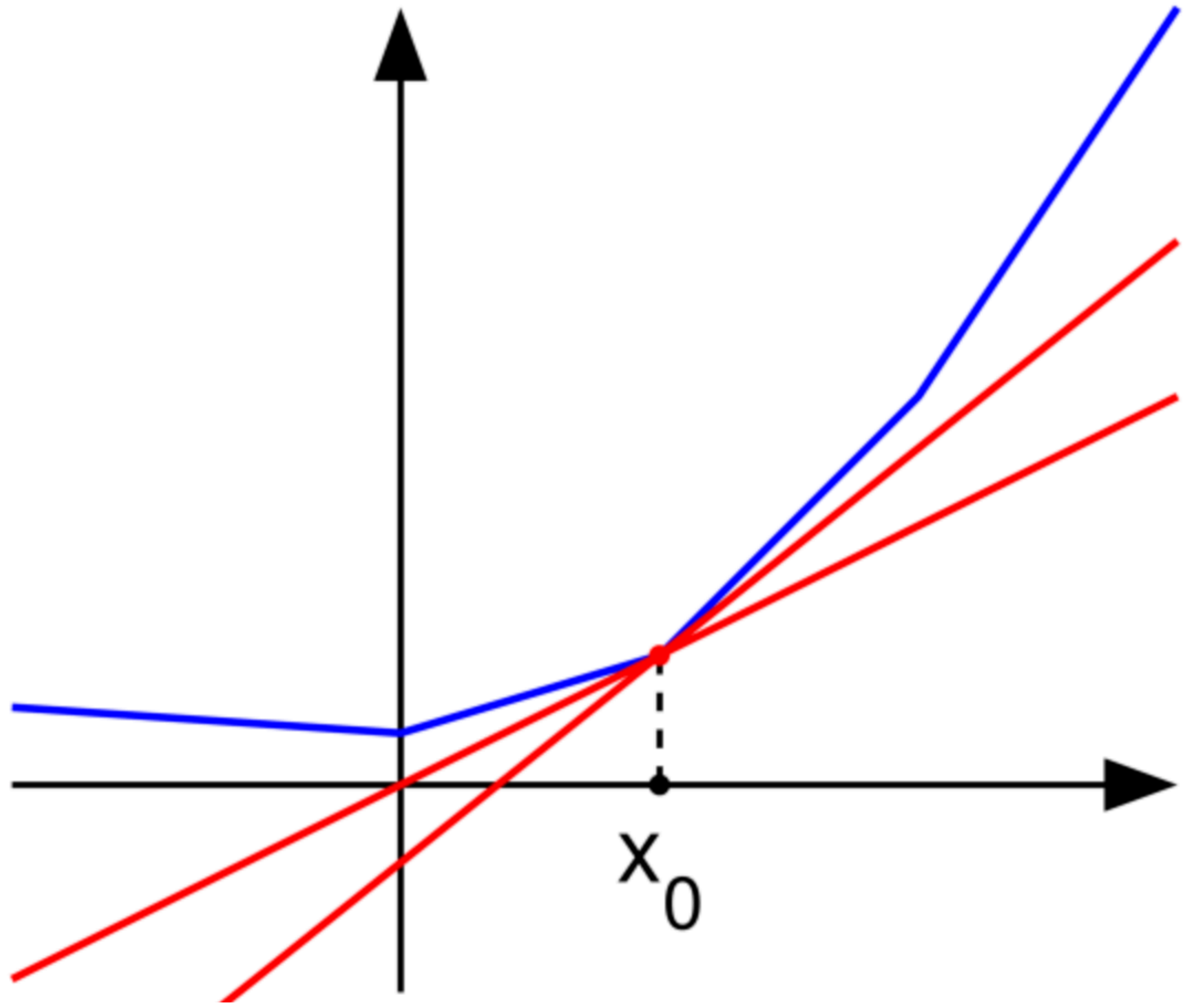
First-order Feedback

Online Projected Sub-gradient Descent (PSGD)

- Algorithm:
  - $\mathbf{w}_1 \in C$ arbitrary.
  - $\mathbf{w}_{t+1} = \Pi_C[\mathbf{w}_t - \eta\, \delta f_t(\mathbf{w}_t)]$, where
    - Π_C is the projection over C.
    - $\delta f_t(\mathbf{w}_t) \in \partial f_t(\mathbf{w}_t)$ (sub-gradient of f_t at $\mathbf{w}_t$).
    - $\eta > 0$ parameter.

Online Projected Sub-gradient Descent (PSGD)

- Algorithm:
  - $\mathbf{w}_1 \in C$ arbitrary.
  - $\mathbf{w}_{t+1} = \Pi_C[\mathbf{w}_t - \eta\, \delta f_t(\mathbf{w}_t)]$, where
    - Π_C is the projection over C.
    - $\delta f_t(\mathbf{w}_t) \in \partial f_t(\mathbf{w}_t)$ (sub-gradient of f_t at $\mathbf{w}_t$).
    - $\eta > 0$ parameter.

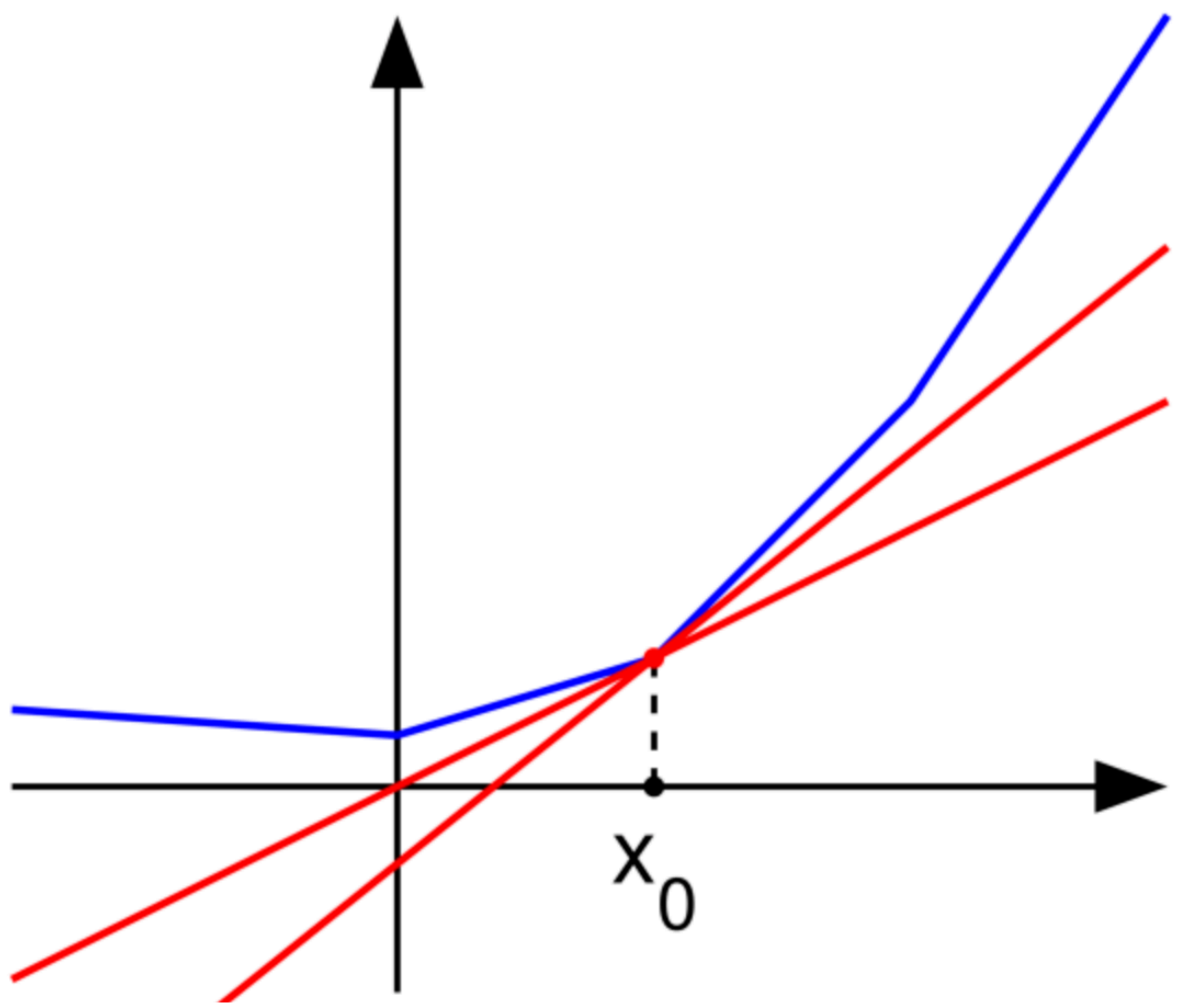


The sub-gradient $\delta f(x_0)$ at $x = x_0$ is the set of all the vector c such that

$$f(x_0) - f(x) \leq c(x_0 - x)\,.$$

Online Projected Sub-gradient Descent (PSGD)

- Algorithm:
  - $\mathbf{w}_1 \in C$ arbitrary.
  - $\mathbf{w}_{t+1} = \Pi_C[\mathbf{w}_t - \eta\, \delta f_t(\mathbf{w}_t)]$, where
    - Π_C is the projection over C.
    - $\delta f_t(\mathbf{w}_t) \in \partial f_t(\mathbf{w}_t)$ (sub-gradient of f_t at $\mathbf{w}_t$).
    - $\eta > 0$ parameter.



The sub-gradient $\delta f(x_0)$ at $x = x_0$ is the set of all the vector c such that

$$f(x_0) - f(x) \leq c(x_0 - x)\,.$$

In almost all cases, the projection function in PSGD is the Euclidean projector

$$\Pi_C(x) = \arg\min_{x' \in C} \|x' - x\|^2.$$

Regret of PSGD [Zinkevich, 2009]

- Assumptions:
  - $\|\mathbf{w}_1 - \mathbf{w}^*\| \leq R$ where $\mathbf{w}^* \in \operatorname{argmin}_{\mathbf{w}\in C} \sum_{t=1}^{T} f_t(\mathbf{w})$.
  - $\|\delta f_t(\mathbf{w}_t)\| \leq G$.
- Theorem: the regret of online projected sub-gradient descent (PSGD) is bounded as follows

$$R_T(\text{PSGD}) \leq \frac{R^2}{2\eta} + \frac{\eta G^2 T}{2}.$$

Choosing η to minimize the bound gives

$$R_T(\text{PSGD}) \leq RG\sqrt{T}.$$

$\|\delta f_t(w_t)\|$

$= \|\nabla f_t(w_t)\|$

$\leq G$

$|f_t(x) - f_t(y)|$

$\leq G(x-y).$

Regret of PSGD [Zinkevich, 2009]

- Assumptions:
  - $\|\mathbf{w}_1 - \mathbf{w}^*\| \leq R$ where $\mathbf{w}^* \in \operatorname{argmin}_{\mathbf{w}\in C} \sum_{t=1}^{T} f_t(\mathbf{w})$.
  - $\|\delta f_t(\mathbf{w}_t)\| \leq G$.
- Theorem: the regret of online projected sub-gradient descent (PSGD) is bounded as follows

$$R_T(\text{PSGD}) \leq \frac{R^2}{2\eta} + \frac{\eta G^2 T}{2}.$$

Choosing η to minimize the bound gives

$$\boxed{R_T(\text{PSGD}) \leq RG\sqrt{T}.}$$

→ The same $O(\sqrt{T})$ bound as FTRL

Proof

- The proof uses the definition of subgradient and the property of projection:

$$
\begin{aligned}
R_T(\text{PSGD}) &= \sum_{t=1}^{T} \left(f_t(\mathbf{w}_t) - f_t(\mathbf{w}^*)\right) \\
&\leq \sum_{t=1}^{T} \delta f_t(\mathbf{w}_t) \cdot (\mathbf{w}_t - \mathbf{w}^*) && \text{(def. of subgrad.)} \\
&= \sum_{t=1}^{T} \frac{1}{2\eta} \left[\|\mathbf{w}_t - \mathbf{w}^*\|^2 \right] + \eta^2 \|\delta f_t(\mathbf{w}_t)\|^2 - \|\mathbf{w}_t - \eta \delta f_t(\mathbf{w}_t) - \mathbf{w}^*\|^2 \Big] \\
&\leq \sum_{t=1}^{T} \frac{1}{2\eta} \left[\|\mathbf{w}_t - \mathbf{w}^*\|^2 \right] + \eta^2 G^2 - \|\mathbf{w}_{t+1} - \mathbf{w}^*\|^2 \Big] && \text{(prop. of proj.)} \\
&\leq \frac{1}{2\eta} \left[\|\mathbf{w}_1 - \mathbf{w}^*\|^2 \right] + \eta^2 G^2 T - \|\mathbf{w}_{T+1} - \mathbf{w}^*\|^2 \Big] \\
&\leq \frac{1}{2\eta} \left[\|\mathbf{w}_1 - \mathbf{w}^*\|^2 \right] + \eta^2 G^2 T \Big] \leq \frac{1}{2\eta} \left[R^2 + \eta^2 G^2 T \right].
\end{aligned}
$$

Application

- Application: $\min_{\mathbf{w} \in C} f(\mathbf{w})$.
  - fixed loss function: $f_t = f$.
  - guarantee for average weight vector:

$$f\left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{w}_t\right) - f(\mathbf{w}^*) \leq \frac{1}{T}\sum_{t=1}^{T} f(\mathbf{w}_t) - f(\mathbf{w}^*)$$

$$= \frac{R_T(\mathcal{A})}{T} = O\left(\frac{1}{\sqrt{T}}\right).$$

$$f(w_T) - f(w^*) \leq \frac{1}{\sqrt{T}}.$$

  - thus, convergence in $O\left(\frac{1}{\epsilon^2}\right)$.

$$\varepsilon = \frac{1}{\sqrt{T}}$$

$$T = \frac{1}{\varepsilon^2}$$

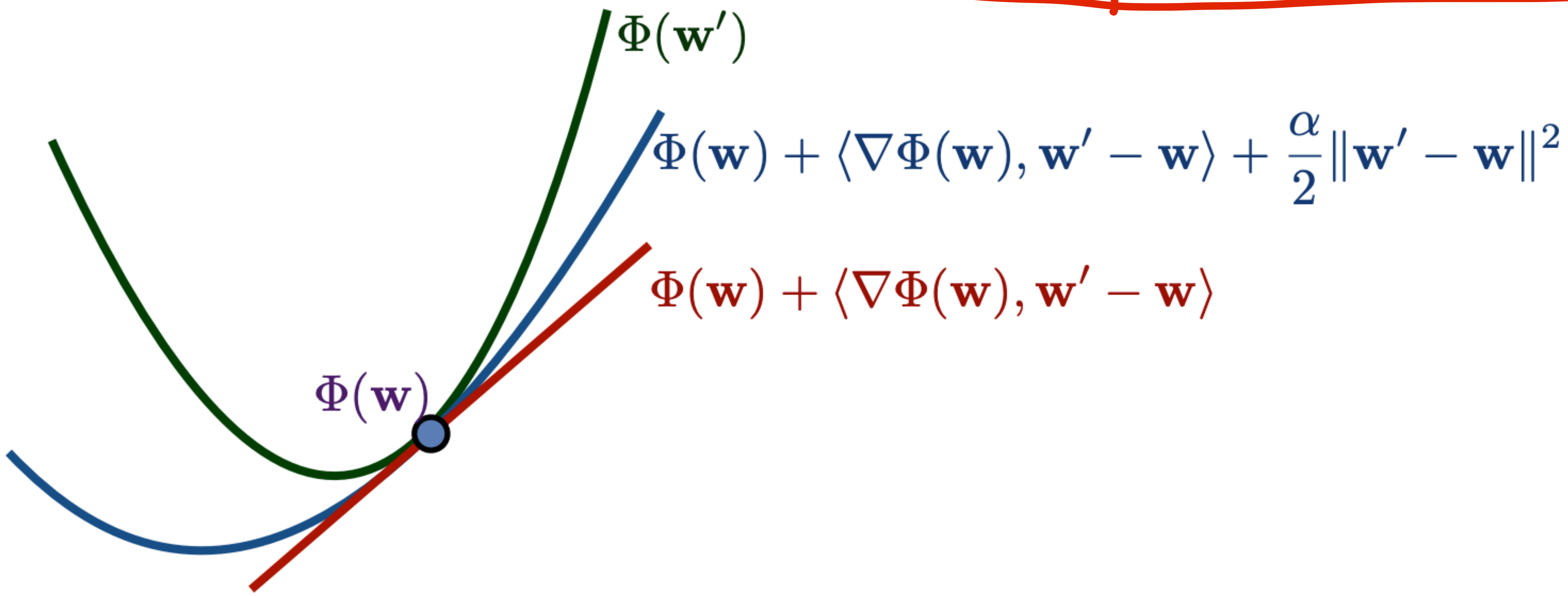
Strongly Convex Loss Functions [Hazan et al., 2007]

- Theorem: assume that the functions f_t are α-strongly convex and $\|\delta f_t(\mathbf{w})\| \leq G$ for all $\mathbf{w}$ and $\delta f_t \in \partial f_t(\mathbf{w})$. Then, the regret of online projected sub-gradient descent (PSGD) with parameter $\eta_{t+1} = \frac{1}{\alpha t}$ is bounded as follows

$$R_T(\text{PSGD}) \leq \frac{G^2}{2\alpha}(1 + \log T).$$

Strong Convexity

- Definition: a convex function Φ defined over a convex set C is α-strongly convex with respect to norm $\|\cdot\|$ if the function $\mathbf{w} \mapsto \Phi(\mathbf{w}) - \frac{\alpha}{2}\|\mathbf{w}\|^2$ is convex or, equivalently,
  - for all $\mathbf{w}, \mathbf{w}'$ in C and $\delta\Phi(\mathbf{w}) \in \partial\Phi(\mathbf{w})$,
    $\Phi(\mathbf{w}') \geq \Phi(\mathbf{w}) + \delta\Phi(\mathbf{w}) \cdot (\mathbf{w}' - \mathbf{w}) + \frac{\alpha}{2}\|\mathbf{w}' - \mathbf{w}\|^2.$



Proof

$$\sum_{t=1}^{T} \eta_t = \sum_{t=1}^{T} \frac{1}{\alpha t} \sim \frac{1}{\alpha} \log T.$$

$$
\begin{aligned}
&R_T(\text{PSGD}) \\
&= \sum_{t=1}^{T} \left(f_t(\mathbf{w}_t) - f_t(\mathbf{w}^*)\right) \\
&\leq \sum_{t=1}^{T} \delta f_t(\mathbf{w}_t) \cdot (\mathbf{w}_t - \mathbf{w}^*) - \frac{\alpha}{2}\|\mathbf{w}_t - \mathbf{w}^*\|^2 && \text{(strong convexity)} \\
&= \sum_{t=1}^{T} \frac{1}{2\eta_{t+1}} \left[\|\mathbf{w}_t - \mathbf{w}^*\|^2 + \eta_{t+1}^2 \|\delta f_t(\mathbf{w}_t)\|^2 - \|\mathbf{w}_t - \eta_{t+1}\delta f_t(\mathbf{w}_t) - \mathbf{w}^*\|^2\right] - \frac{\alpha}{2}\|\mathbf{w}_t - \mathbf{w}^*\|^2 \\
&\leq \sum_{t=1}^{T} \frac{1}{2\eta_{t+1}} \left[\|\mathbf{w}_t - \mathbf{w}^*\|^2 + \eta_{t+1}^2 G^2 - \|\mathbf{w}_{t+1} - \mathbf{w}^*\|^2\right] - \frac{\alpha}{2}\|\mathbf{w}_t - \mathbf{w}^*\|^2 && \text{(prop. of proj.)} \\
&\leq \frac{\alpha}{2} \sum_{t=1}^{T} \left[(t-1)\|\mathbf{w}_t - \mathbf{w}^*\|^2 - t\|\mathbf{w}_{t+1} - \mathbf{w}^*\|^2\right] + \frac{G^2}{2\alpha} \sum_{t=1}^{T} \frac{1}{t} && \text{(def. of } \eta_{t+1}\text{)} \\
&= \frac{\alpha}{2}\left[-T\|\mathbf{w}_{T+1} - \mathbf{w}^*\|^2\right] + \frac{G^2}{2\alpha} \sum_{t=1}^{T} \frac{1}{t} \leq \frac{G^2}{2\alpha} \sum_{t=1}^{T} \frac{1}{t} \leq \frac{G^2}{2\alpha}(1 + \log T).
\end{aligned}
$$

Exponentiated Gradient (EG) [Kivinen and Warmuth, 1997]

- Convex set: simplex $C = \{\mathbf{w} \in \mathbb{R}^N : \mathbf{w} \geq 0 \wedge \|\mathbf{w}\|_1 = 1\}$.
- Algorithm:
  - $\mathbf{w}_1 = (\frac{1}{N}, \ldots, \frac{1}{N})^\top$.
  - $\mathbf{w}_{t+1,i} = \dfrac{\mathbf{w}_{t,i} \exp(-\eta \, [\delta f_t(\mathbf{w}_t)]_i)}{Z_t}$ where

    $Z_t = \sum_{i=1}^N \mathbf{w}_{t,i} e^{-\eta [\delta f_t(\mathbf{w}_t)]_i}$.

Exponentiated Gradient (EG) [Kivinen and Warmuth, 1997]

- Convex set: simplex $C = \{\mathbf{w} \in \mathbb{R}^N : \mathbf{w} \geq 0 \wedge \|\mathbf{w}\|_1 = 1\}$.
- Algorithm:
  - $\mathbf{w}_1 = (\frac{1}{N}, \ldots, \frac{1}{N})^\top$.
  - $\mathbf{w}_{t+1,i} = \dfrac{\mathbf{w}_{t,i} \exp(-\eta\, [\delta f_t(\mathbf{w}_t)]_i)}{Z_t}$ where

$$Z_t = \sum_{i=1}^N \mathbf{w}_{t,i} e^{-\eta[\delta f_t(\mathbf{w}_t)]_i}.$$

$f_t(w_t) = (l_t)\, w_t$

$\delta f_t(w_t) = (l_{t,i})_{i=1}^N$

When the loss function $f_t(w_t) = l_t \cdot w_t$ takes the linear form, it is multiplicative weight update algorithm.

Regret of EG

- Assumption:
  - $\|\delta f_t(\mathbf{w}_t)\|_\infty \leq G_\infty$.
- Theorem: the regret of the Exponentiated Gradient (EG) algorithm is bounded as follows

$$R_T(\text{EG}) \leq \frac{\log N}{\eta} + \frac{\eta G_\infty^2 T}{2}.$$

Choosing η to minimize the bound gives

$$\boxed{R_T(\text{EG}) \leq 2G_\infty \sqrt{T \log N}.}$$

Regret of EG

- Assumption:
  - $\|\delta f_t(\mathbf{w}_t)\|_\infty \leq G_\infty$.
- Theorem: the regret of the Exponentiated Gradient (EG) algorithm is bounded as follows

$$R_T(\text{EG}) \leq \frac{\log N}{\eta} + \frac{\eta G_\infty^2 T}{2}.$$

Choosing η to minimize the bound gives

$$R_T(\text{EG}) \leq 2G_\infty\sqrt{T\log N}.$$

$$R_T(\text{PSGD}) \leq RG\sqrt{T}.$$

$G = \|\delta f_t(w_t)\|_2^2 \approx O(\sqrt{N})$

Comparison between EG & PSGD

In EG, $G_\infty \sim O(1)$,
which yields the regret $O(\sqrt{T\log N})$

In PSGD, $G \sim O(\sqrt{N})$,
which yields the regret $O(\sqrt{TN})$

Proof

$$\mathbb{E}\, e^{-\eta(x - \mathbb{E}x)}$$

- Potential: $\Phi_t = D(\mathbf{w}^* \parallel \mathbf{w}_t) = \sum_{i=1}^{N} \mathbf{w}_i^* \log \frac{\mathbf{w}_i^*}{\mathbf{w}_{t,i}}$.
- $\Phi_{t+1} - \Phi_t = \sum_{i=1}^{N} \mathbf{w}_i^* \log \frac{\mathbf{w}_{t,i}}{\mathbf{w}_{t+1,i}}$

$$= \sum_{i=1}^{N} \mathbf{w}_i^* \left[\log Z_t + \eta [\delta f_t(\mathbf{w}_t)]_i \right] = \log Z_t + \eta \mathbf{w}^* \cdot \delta f_t(\mathbf{w}_t).$$

- $\log Z_t = \log \left[\sum_{i=1}^{N} \mathbf{w}_{t,i} e^{-\eta [\delta f_t(\mathbf{w}_t)]_i} \right]$

$$\rightarrow \mathbb{E}_{w_t}\, e^{-\eta[\delta f_t(w_t)]_i}$$

$$= \log \mathop{\mathrm{E}}_{i \sim \mathbf{w}_t} \left[e^{-\eta [\delta f_t(\mathbf{w}_t)]_i} \right]$$

$$= \log \mathop{\mathrm{E}}_{i \sim \mathbf{w}_t} \left[e^{-\eta \left([\delta f_t(\mathbf{w}_t)]_i - \mathrm{E}\left[[\delta f_t(\mathbf{w}_t)]_i \right] \right) - \eta\, \mathrm{E}\left[[\delta f_t(\mathbf{w}_t)]_i \right]} \right]$$

$$\leq \eta^2 \frac{4 G_\infty^2}{8} - \eta \mathbf{w}_t \cdot \delta f_t(\mathbf{w}_t) \qquad \text{(Hoeffding's ineq.).}$$

Proof

- Potential: $\Phi_t = D(\mathbf{w}^* \parallel \mathbf{w}_t) = \sum_{i=1}^{N} \mathbf{w}_i^* \log \frac{\mathbf{w}_i^*}{\mathbf{w}_{t,i}}$.
- $$\Phi_{t+1} - \Phi_t = \sum_{i=1}^{N} \mathbf{w}_i^* \log \frac{\mathbf{w}_{t,i}}{\mathbf{w}_{t+1,i}}$$
$$= \sum_{i=1}^{N} \mathbf{w}_i^* \left[\log Z_t + \eta[\delta f_t(\mathbf{w}_t)]_i \right] = \log Z_t + \eta \mathbf{w}^* \cdot \delta f_t(\mathbf{w}_t).$$
- $$\log Z_t = \log \left[\sum_{i=1}^{N} \mathbf{w}_{t,i} e^{-\eta[\delta f_t(\mathbf{w}_t)]_i} \right]$$
$$= \log \mathop{\mathrm{E}}_{i \sim \mathbf{w}_t} \left[e^{-\eta[\delta f_t(\mathbf{w}_t)]_i} \right]$$
$$= \log \mathop{\mathrm{E}}_{i \sim \mathbf{w}_t} \left[e^{-\eta\left([\delta f_t(\mathbf{w}_t)]_i - \mathrm{E}\left[[\delta f_t(\mathbf{w}_t)]_i\right]\right) - \eta \mathrm{E}\left[[\delta f_t(\mathbf{w}_t)]_i\right]} \right]$$
$$\leq \eta^2 \frac{4G_\infty^2}{8} - \eta \mathbf{w}_t \cdot \delta f_t(\mathbf{w}_t) \quad \text{(Hoeffding's ineq.)}.$$

Heoffding's Ineq.

Suppose X is a real random variable such that $\mathrm{P}(X \in [a,b]) = 1$. Then
$$\mathrm{E}\left[e^{s(X - \mathrm{E}[X])}\right] \leq \exp\left(\tfrac{1}{8} s^2 (b-a)^2\right).$$

$s = -\eta$
$x = [\delta f_t(W_t)]_i$
$b = G_\infty \quad a = -G_\infty$

Proof

- Combining equality and inequality:

$$\Phi_{t+1} - \Phi_t \le \frac{\eta^2 G_\infty^2}{2} - \eta(\mathbf{w}^* - \mathbf{w}_t) \cdot \delta f_t(\mathbf{w}_t)$$

$$\Leftrightarrow \eta(\mathbf{w}^* - \mathbf{w}_t) \cdot \delta f_t(\mathbf{w}_t) \le \frac{\eta^2 G_\infty^2}{2} + (\Phi_t - \Phi_{t+1})$$

$$\Rightarrow \sum_{t=1}^{T} (\mathbf{w}^* - \mathbf{w}_t) \cdot \delta f_t(\mathbf{w}_t) \le \frac{\eta^2 G_\infty^2 T}{2} + \frac{\Phi_1 - \Phi_{T+1}}{\eta}$$

$$\Rightarrow \sum_{t=1}^{T} (\mathbf{w}^* - \mathbf{w}_t) \cdot \delta f_t(\mathbf{w}_t) \le \frac{\eta^2 G_\infty^2 T}{2} + \frac{\Phi_1}{\eta}. \qquad \text{(Rel. Ent. non-neg.)}$$

- $$R_T(\mathrm{EG}) = \sum_{t=1}^{T} \big(f_t(\mathbf{w}_t) - f_t(\mathbf{w}^*)\big)$$

$$\le \sum_{t=1}^{T} \delta f_t(\mathbf{w}_t) \cdot (\mathbf{w}_t - \mathbf{w}^*)$$

$$\le \frac{\eta G_\infty^2 T}{2} + \frac{\Phi_1}{\eta} = \frac{\eta G_\infty^2 T}{2} + \frac{D(\mathbf{w}^* \parallel \mathbf{w}_1)}{\eta} \le \frac{\eta G_\infty^2 T}{2} + \frac{\log N}{\eta}.$$

$\eta = \sqrt{\frac{1}{T \log N}}$

Online Mirror Descent (OMD)

- PSGD and EG both special instances of a more general algorithm: Mirror Descent.

- Mirror Descent is based on Bregman divergence:
  - PSGD: $B(\mathbf{w} \parallel \mathbf{w}') = \frac{1}{2}\|\mathbf{w} - \mathbf{w}'\|_2^2$.
  - EG: unnormalized relative entropy;

$$B(\mathbf{w} \parallel \mathbf{w}') = \sum_{i=1}^{N} \left[w_i \log\left[\frac{w_i}{w_i'}\right] - w_i + w_i' \right].$$

Online Mirror Descent (OMD)

- PSGD and EG both special instances of a more general algorithm: Mirror Descent.

- Mirror Descent is based on Bregman divergence:
  - PSGD: $B(\mathbf{w} \,\|\, \mathbf{w}') = \frac{1}{2}\|\mathbf{w} - \mathbf{w}'\|_2^2$.
  - EG: unnormalized relative entropy;

$$B(\mathbf{w} \,\|\, \mathbf{w}') = \sum_{i=1}^{N} \left[w_i \log\left[\frac{w_i}{w_i'}\right] - w_i + w_i' \right].$$

Will dive into OMD later. Let's first look at the zeroth-order feedback setting.

Zeroth-order Feedback

Adversarial Bandit Problem

- The optimization space $\mathscr{X} = [n]$ is discrete, and the loss function f_t becomes loss vector $l_t = (l_{t,1}, \cdots, l_{t,n})$.
- At each step t , the adversary picks a loss vector l_t .
- The optimizer draws an action $I_t \sim p_t, p_t \in \Delta_n$ is a probability distribution over $[n]$.
- The optimizer only observes the loss value l_{t,I_t}.

Adversarial Bandit Problem

- The optimization space $\mathscr{X} = [n]$ is discrete, and the loss function f_t becomes loss vector $l_t = (l_{t,1}, \cdots, l_{t,n})$.
- At each step t , the adversary picks a loss vector l_t .
- The optimizer draws an action $I_t \sim p_t, p_t \in \Delta_n$ is a probability distribution over $[n]$.
- The optimizer only observes the loss value l_{t,I_t}.

Remarks:

1. This is an online convex optimization problem with linear loss function under the zeroth order feedback setting.
2. If the optimizer can observe the whole loss vector l_t, this is exactly the online convex optimization setting in EG. Just let $C = \Delta_n$, and $w_t = p_t$, $f_t(w_t) = l_t \cdot w_t$.
3. To achieve no-regret, we need an updated version of EG.

Exponential-weight Algorithm for Exploration and Exploitation (EXP3)

Algorithm 1 EXP3

$w_1 = (1, \ldots, 1)$
for $t = 1$ to T **do**
  Define $p_t = \frac{w_t}{\|w_t\|}$
  Draw $I_t \sim p_t$
  Observe $\ell_{t,I_t} \in [0, 1]$
  for $i = 1$ to d **do**
    if $i \neq I_t$ **then**
      $w_{t+1,i} = w_{t,i}$
    else
      $w_{t+1,i} = w_{t,i} e^{-\eta \frac{\ell_{t,i}}{p_{t,i}}}$
    end if
  end for
end for

Exponential-weight Algorithm for Exploration and Exploitation (EXP3)

Algorithm 1 EXP3

$w_1 = (1, \dots, 1)$
for $t = 1$ to T **do**
  Define $p_t = \frac{w_t}{\|w_t\|}$
  Draw $I_t \sim p_t$
  Observe $\ell_{t,I_t} \in [0, 1]$
  for $i = 1$ to d **do**
    if $i \neq I_t$ **then**
      $w_{t+1,i} = w_{t,i}$
    else
      $w_{t+1,i} = w_{t,i} e^{-\eta \frac{\ell_{t,i}}{p_{t,i}}}$
    end if
  end for
end for

Replace the loss vector

$$l_t = (l_{t,1}, \cdots, l_{t,n})$$

by the unbiased estimator

$$\hat{l}_t = (0, \cdots, 0, \frac{l_{t,i}}{p_{t,i}}, 0, \cdots, 0)$$

Exponential-weight Algorithm for Exploration and Exploitation (EXP3)

Algorithm 1 EXP3

$w_1 = (1, \ldots, 1)$
for $t = 1$ to T **do**
  Define $p_t = \frac{w_t}{\|w_t\|}$
  Draw $I_t \sim p_t$
  Observe $\ell_{t,I_t} \in [0, 1]$
  for $i = 1$ to d **do**
    if $i \neq I_t$ **then**
      $w_{t+1,i} = w_{t,i}$
    else
      $w_{t+1,i} = w_{t,i} e^{-\eta \frac{\ell_{t,i}}{p_{t,i}}}$
    end if
  end for
end for

Replace the loss vector

$$l_t = (l_{t,1}, \cdots, l_{t,n})$$

by the unbiased estimator

$$\hat{l}_t = (0, \cdots, 0, \frac{l_{t,i}}{p_{t,i}}, 0, \cdots, 0)$$

Theorem

$$\mathbb{E}\left[\text{Regret}_T(\text{EXP3})\right] \leq \sqrt{2TN \ln N}$$

Proof

Let $\Phi_t = \frac{1}{\eta} \ln(\sum_{a=1}^{N} \exp(-\eta \sum_{s=1}^{t} l_t(a)))$, we have

$$\Phi_T - \Phi_0 = \sum_{t=1}^{T} \Phi_t - \Phi_{t-1} = \sum_{t=1}^{T} \frac{1}{\eta} \ln(\sum_{a=1}^{N} w_t(a) \exp(-\eta l_t(a))).$$

Therefore, $\Phi_T - \Phi_0 = \sum_{t=1}^{T} \frac{1}{\eta} \ln\left(\sum_{a=1}^{N} w_t(a) \exp(-\eta l_t(a))\right)$

$$\leq \sum_{t=1}^{T} \frac{1}{\eta} \ln\left(\sum_{a=1}^{N} w_t(a)\left[1 - \eta l_t(a) + \frac{1}{2}\eta^2 l_t(a)^2\right]\right)$$

$$\leq \sum_{t=1}^{T} \frac{1}{\eta} \ln\left(\left[1 - \eta \sum_{a=1}^{N} w_t(a) l_t(a) + \frac{1}{2}\eta^2 \sum_{a=1}^{N} w_t(a) l_t(a)^2\right]\right)$$

$$\leq \sum_{t=1}^{T} \frac{1}{\eta}\left[- \eta \sum_{a=1}^{N} w_t(a) l_t(a) + \frac{1}{2}\eta^2 \sum_{a=1}^{N} w_t(a) l_t(a)^2\right]$$

$$\leq \sum_{t=1}^{T}\left[- \boldsymbol{w}_t \cdot \boldsymbol{l}_t + \eta \sum_{a=1}^{N} w_t(a) l_t(a)^2\right]$$

Proof

Let $\Phi_t = \frac{1}{\eta}\ln(\sum_{a=1}^{N}\exp(-\eta\sum_{s=1}^{t} l_t(a)))$, we have

$$\Phi_T - \Phi_0 = \sum_{t=1}^{T}\Phi_t - \Phi_{t-1} = \sum_{t=1}^{T}\frac{1}{\eta}\ln(\sum_{a=1}^{N} w_t(a)\exp(-\eta l_t(a)))\,.$$

Therefore, $\Phi_T - \Phi_0 = \sum_{t=1}^{T}\frac{1}{\eta}\ln\left(\sum_{a=1}^{N} w_t(a)\exp(-\eta l_t(a))\right)$

$$\leq \sum_{t=1}^{T}\frac{1}{\eta}\ln\left(\sum_{a=1}^{N} w_t(a)\left[1-\eta l_t(a)+\frac{1}{2}\eta^2 l_t(a)^2\right]\right)$$

$$\leq \sum_{t=1}^{T}\frac{1}{\eta}\ln\left(\left[1-\eta\sum_{a=1}^{N} w_t(a)l_t(a)+\frac{1}{2}\eta^2\sum_{a=1}^{N} w_t(a)l_t(a)^2\right]\right)$$

$$\leq \sum_{t=1}^{T}\frac{1}{\eta}\left[-\eta\sum_{a=1}^{N} w_t(a)l_t(a)+\frac{1}{2}\eta^2\sum_{a=1}^{N} w_t(a)l_t(a)^2\right]$$

$$\leq \sum_{t=1}^{T}\left[-\boldsymbol{w}_t\cdot\boldsymbol{l}_t+\eta\sum_{a=1}^{N} w_t(a)l_t(a)^2\right]$$

For any sub-Gaussian r.v. X with zero mean,

$$\mathbb{E}[\exp(sX)] \leq \exp\left(\frac{s^2\mathbb{E}X^2}{2}\right).$$

Proof

Note that $\Phi_0 = \frac{1}{\eta} \ln(\sum_{a=1}^{N} 1) = \frac{1}{\eta} \ln N,$ we have

$$\begin{aligned}
\Phi_T - \Phi_0 &\leq -\sum_{t=1}^{T} \boldsymbol{w}_t \cdot \boldsymbol{l}_t + \eta \sum_{t=1}^{T} \sum_{a=1}^{N} w_t(a) l_t(a)^2 \\
\Phi_T - \frac{1}{\eta} \ln N &\leq \\
\Phi_T + \sum_{t=1}^{T} \boldsymbol{w}_t \cdot \boldsymbol{l}_t &\leq \frac{1}{\eta} \ln N + \eta \sum_{t=1}^{T} \sum_{a=1}^{N} w_t(a) l_t(a)^2
\end{aligned}$$

Also note that $\Phi_T \geq -\sum_{t=1}^{T} l_t(a)$, we have

$$\begin{aligned}
\sum_{t=1}^{T} \boldsymbol{w}_t \cdot \boldsymbol{l}_t + \Phi_T &\leq \frac{\ln N}{\eta} + \eta \sum_{t=1}^{t} \sum_{a=1}^{N} w_t(a) l_t(a)^2 \\
\sum_{t=1}^{T} \boldsymbol{w}_t \cdot \boldsymbol{l}_t - L_T(a) &\leq
\end{aligned}$$

Proof

Replace l_t with $\hat{l}_t$, we obtain

$$
\begin{aligned}
\mathbb{E}\left[\sum_{t=1}^{T} \boldsymbol{w}_t \cdot \hat{\boldsymbol{l}}_t - \min_a L_T(a)\right] &\le \frac{\ln N}{\eta} + \eta \sum_{t=1}^{T}\sum_{a=1}^{N} \mathbb{E}\left[w_t(a)\cdot \hat{l}_t(a)^2\right] \\
&\le \frac{\ln N}{\eta} + \eta \sum_{t=1}^{T}\sum_{a=1}^{N} \mathbb{E}\left[w_t(a)\right] \cdot \frac{1}{\Pr(a)} l_i(a)^2 \\
&\le \frac{\ln N}{\eta} + \eta \sum_{t=1}^{T}\sum_{a=1}^{N} l_i(a)^2 \\
&\le \frac{\ln N}{\eta} + \eta \sum_{t=1}^{T}\sum_{a=1}^{N} 1 \\
&\le \frac{\ln N}{\eta} + \eta T N
\end{aligned}
$$

Set $\eta = \sqrt{\frac{\ln N}{TN}}$, we have $\mathbb{E}\left[\text{Regret}_T(\text{EXP3})\right] \le \sqrt{2TN\ln N}$

$\mathrm{Var}[\hat{l}_t(a)] = \dots$

$\hat{l}_t(a) \begin{cases} 0 \\ \frac{l_t(a)}{P_t(a)} \end{cases}$

Variants of EXP3

- The regret bound for EXP3 only holds in expectation (Pseudo Regret). To derive the high-probability bound for the true regret, we have two variants of EXP3:

1. EXP3-P [Auer, 2001]

   Uniform exploration: $\tilde{l}_t(a) = l_t(a) + \frac{\beta}{p_t(a)}, \beta \sim \sqrt{\frac{\log NT/\delta}{NT}}$

   Biased loss estimation: $p_t(a) = (1 - \varepsilon)w_t(a) + \frac{\varepsilon}{N}, \varepsilon \sim \sqrt{\frac{N \log N}{T}}$

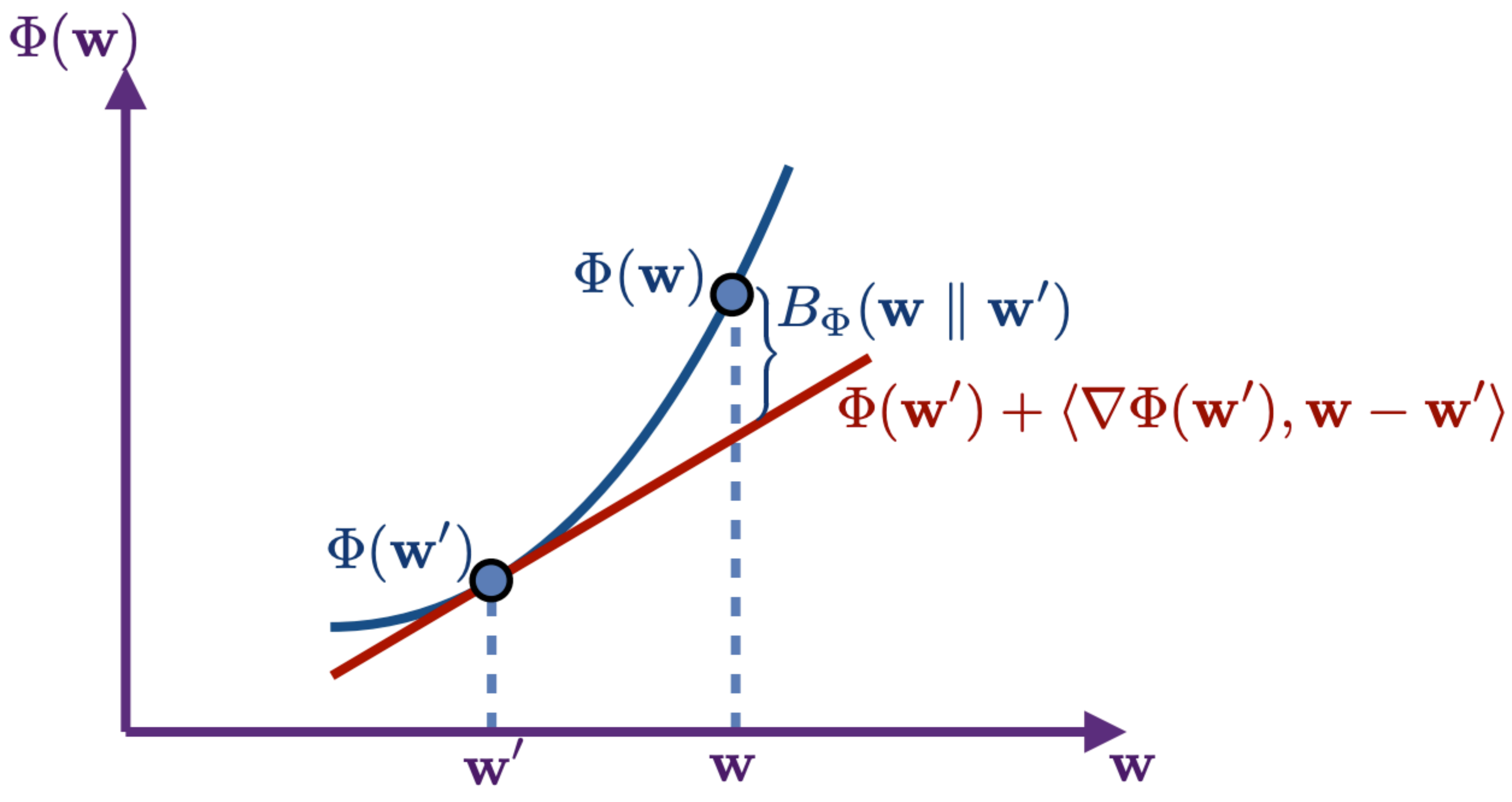
2. EXP3 with Implicit Exploration (EXP-IX) [Neu, 2015]

   Biased loss estimation: $\tilde{l}_t(a) = \frac{l_t(a)}{p_t(a) + \varepsilon_t} \mathbb{1}_{\{I_t = a\}}, \varepsilon_t \sim \sqrt{\frac{\log N}{Nt}}.$

OMD revisited

- Bregman Divergence
  - Definition: Φ convex differentiable over open convex set C.
    The Bregman divergence associated to Φ is defined by

$$B_\Phi(\mathbf{w} \parallel \mathbf{w}') = \Phi(\mathbf{w}) - \Phi(\mathbf{w}') - \langle \nabla\Phi(\mathbf{w}'), \mathbf{w} - \mathbf{w}' \rangle.$$

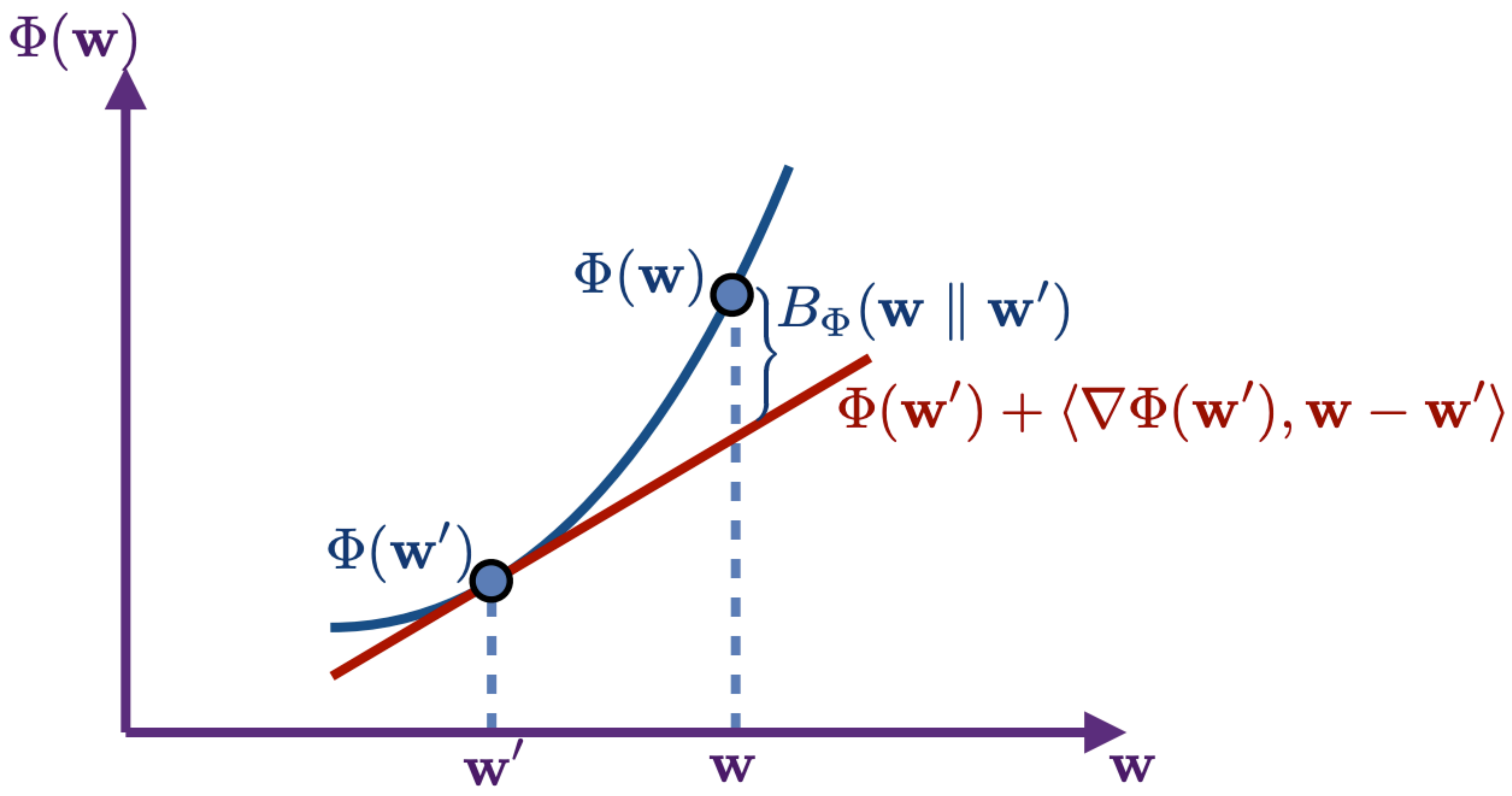


OMD revisited

- Bregman Divergence

Definition: Φ convex differentiable over open convex set C. The Bregman divergence associated to Φ is defined by

$$B_\Phi(\mathbf{w} \parallel \mathbf{w}') = \Phi(\mathbf{w}) - \Phi(\mathbf{w}') - \langle \nabla\Phi(\mathbf{w}'), \mathbf{w} - \mathbf{w}' \rangle.$$



Given any convex function Φ on C, B_Φ is a 'metric' associated with Φ.

Properties of Bregman Divergence

- Proposition: the following properties hold for a Bregman divergence.
  - non-negativity: $\forall \mathbf{w}, \mathbf{w}' \in C, \ B_\Phi(\mathbf{w} \parallel \mathbf{w}') \geq 0.$
  - linearity: $B_{\alpha\Phi+\beta\Psi} = \alpha B_\Phi + \beta B_\Psi$.
  - projection: for any closed convex set $K \subseteq \overline{C}$, the projection of B_Φ-projection of $\mathbf{w}'$ over K is unique:

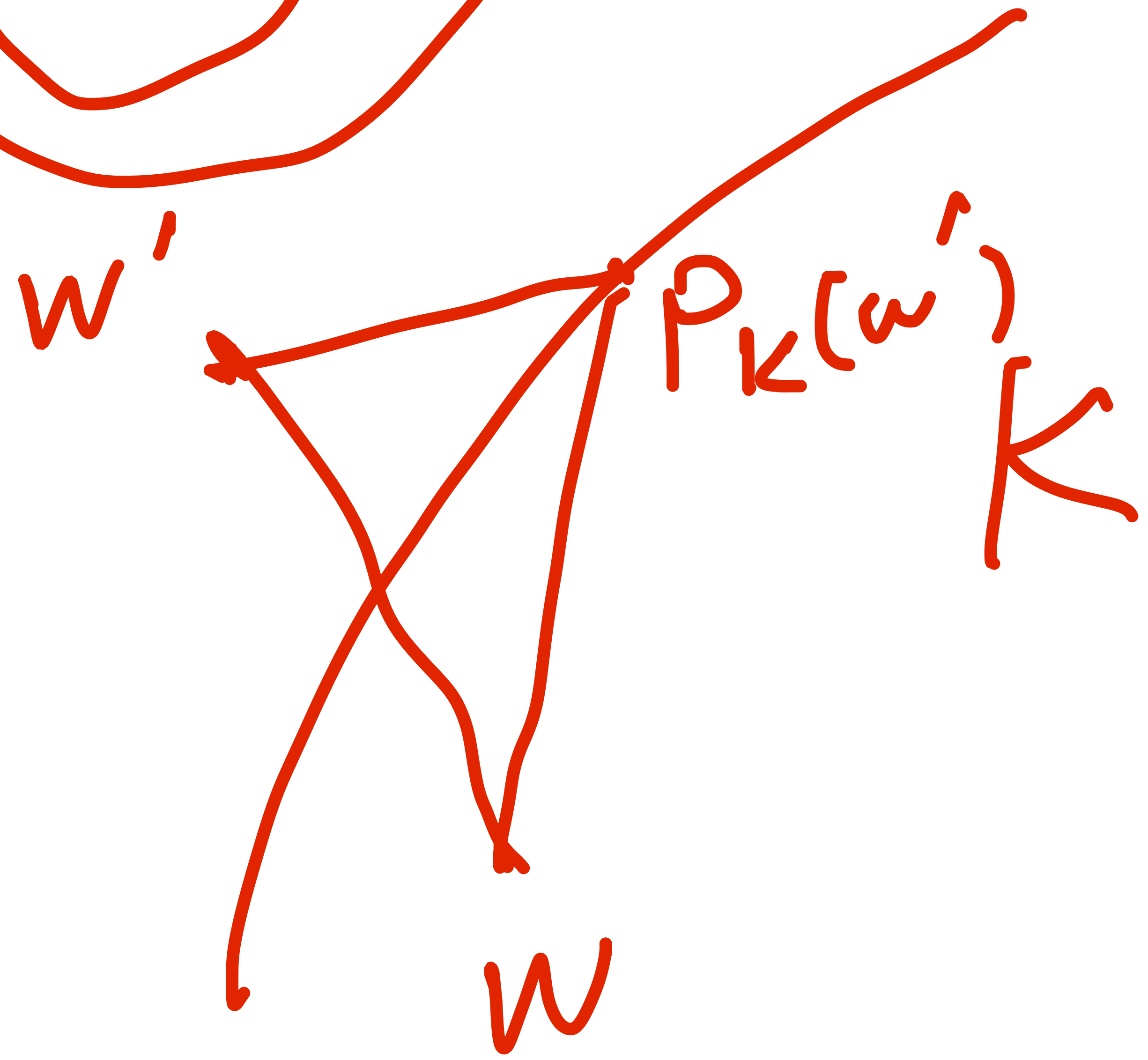
$$P_K(\mathbf{w}') = \underset{\mathbf{w}\in K}{\operatorname{argmin}} B_F(\mathbf{w} \parallel \mathbf{w}').$$

  - Triangular identity:

$$(\nabla\Phi(\mathbf{w}) - \nabla\Phi(\mathbf{v})) \cdot (\mathbf{w} - \mathbf{u}) = B(\mathbf{u} \parallel \mathbf{w}) + B(\mathbf{w} \parallel \mathbf{v}) - B(\mathbf{u} \parallel \mathbf{v}).$$

  - Pythagorean theorem:

$$B_\Phi(\mathbf{w} \parallel \mathbf{w}') \geq B_\Phi(\mathbf{w} \parallel P_K(\mathbf{w}')) + B_\Phi(P_K(\mathbf{w}') \parallel \mathbf{w}').$$



Legendre Type Functions [Rockafellar, 1970]

- **Definition**: a real-valued function Φ defined over a non-empty open convex set C is said to be of Legendre type if it is proper closed convex and differentiable over C and if one of the following equivalent conditions holds:
  - $\nabla\Phi$ is one-to-one mapping from C to $\nabla\Phi(C)$.
  - $\lim_{\mathbf{w}\to\partial C} \|\nabla\Phi(\mathbf{w})\| = +\infty$.

1. $\Phi(x) = \frac{1}{2}\|x\|^2$ $\nabla\Phi(x) = x$ $\nabla^{-1}\Phi(x) = x$

2. $\Phi(x) = \sum x_i \log x_i$, $\nabla\Phi(x) = \log x + 1$

$x \in \Delta_n$ $\nabla^{-1}\Phi(x) = \exp(x-1)$

OMD Algorithm [Nemirovski and Yudin, 1983]

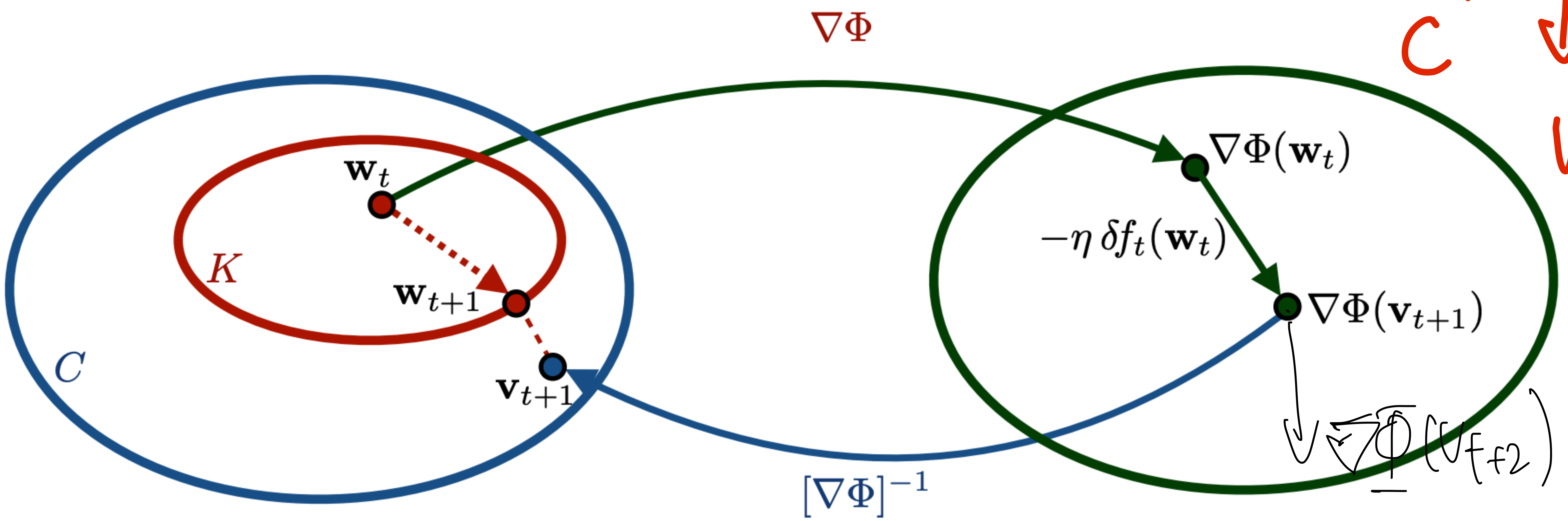
Mirror-Descent(Φ)

$\mathbf{w}_1 \leftarrow \mathrm{argmin}_{\mathbf{w} \in K \cap C}\, \Phi(\mathbf{w})$

for $t \leftarrow 1$ **to** T **do**

$\quad \mathbf{v}_{t+1} \leftarrow [\nabla \Phi]^{-1}\big(\nabla \Phi(\mathbf{w}_t) - \eta\, \delta f_t(\mathbf{w}_t)\big)$

$\quad \mathbf{w}_{t+1} \leftarrow \mathrm{argmin}_{\mathbf{w} \in K \cap C}\, B(\mathbf{w} \parallel \mathbf{v}_{t+1})$



Regret of OMD

- **Theorem**: let C be a non-empty open convex set and $K \subset \overline{C}$ a compact convex set. Assume that $\Phi\colon C \to \mathbb{R}$ is of Legendre type and α-strongly convex with respect to $\|\cdot\|$ and f_ts convex and G_*-Lipschitz with respect to $\|\cdot\|$. Then, the regret of Mirror Descent can be bounded as follows:

$$R_T(\mathrm{MD}) \leq \frac{B(\mathbf{w}^* \parallel \mathbf{w}_1)}{\eta} + \frac{\eta G_*^2 T}{2\alpha}.$$

Choosing η to minimize the bound gives

$$R_T(\mathrm{MD}) \leq D_\Phi G_* \sqrt{\frac{2T}{\alpha}},$$

with $B(\mathbf{w}^* \parallel \mathbf{w}_1) \leq D_\Phi^2$.

$\frac{1}{2}\|x\|^2$

$\|x\|_\infty$

$\Phi(x) = \sum x_i \log x_i$

$\leq \Phi(x') + \nabla\Phi(x)(x'-x) + \frac{\alpha}{2}\|x\|_\infty^2$

Proof

$$
\begin{aligned}
&R_T(\mathrm{MD}) \\
&= \sum_{t=1}^{T} \left(f_t(\mathbf{w}_t) - f_t(\mathbf{w}^*) \right) \\
&\leq \sum_{t=1}^{T} \delta f_t(\mathbf{w}_t) \cdot (\mathbf{w}_t - \mathbf{w}^*) && \text{(def. of subgrad.)} \\
&= \frac{1}{\eta} \sum_{t=1}^{T} [\nabla\Phi(\mathbf{w}_t) - \nabla\Phi(\mathbf{v}_{t+1})] \cdot (\mathbf{w}_t - \mathbf{w}^*) && \text{(def. of } \mathbf{v}_t\text{)} \\
&= \frac{1}{\eta} \sum_{t=1}^{T} \left[B(\mathbf{w}^* \parallel \mathbf{w}_t) - B(\mathbf{w}^* \parallel \mathbf{v}_{t+1}) + B(\mathbf{w}_t \parallel \mathbf{v}_{t+1}) \right] && \text{(Breg. div. Identity)} \\
&\leq \frac{1}{\eta} \sum_{t=1}^{T} \left[B(\mathbf{w}^* \parallel \mathbf{w}_t) - B(\mathbf{w}^* \parallel \mathbf{w}_{t+1}) - B(\mathbf{w}_{t+1} \parallel \mathbf{v}_{t+1}) + B(\mathbf{w}_t \parallel \mathbf{v}_{t+1}) \right] && \text{(Pythagorean ineq.)} \\
&= \frac{1}{\eta} \left[B(\mathbf{w}^* \parallel \mathbf{w}_1) - B(\mathbf{w}^* \parallel \mathbf{w}_{T+1}) \right] + \frac{1}{\eta} \sum_{t=1}^{T} \left[-B(\mathbf{w}_{t+1} \parallel \mathbf{v}_{t+1}) + B(\mathbf{w}_t \parallel \mathbf{v}_{t+1}) \right] \\
&\leq \frac{B(\mathbf{w}^* \parallel \mathbf{w}_1)}{\eta} + \frac{1}{\eta} \sum_{t=1}^{T} \left[B(\mathbf{w}_t \parallel \mathbf{v}_{t+1}) - B(\mathbf{w}_{t+1} \parallel \mathbf{v}_{t+1}) \right].
\end{aligned}
$$

Proof

$$
\begin{aligned}
&\Big[B(\mathbf{w}_t \parallel \mathbf{v}_{t+1}) - B(\mathbf{w}_{t+1} \parallel \mathbf{v}_{t+1})\Big] \\
&= \Phi(\mathbf{w}_t) - \Phi(\mathbf{w}_{t+1}) - \nabla\Phi(\mathbf{v}_{t+1}) \cdot (\mathbf{w}_t - \mathbf{w}_{t+1}) \\
&\leq \big(\nabla\Phi(\mathbf{w}_t) - \nabla\Phi(\mathbf{v}_{t+1})\big) \cdot (\mathbf{w}_t - \mathbf{w}_{t+1}) - \frac{\alpha}{2}\|\mathbf{w}_t - \mathbf{w}_{t+1}\|^2 && (\alpha\text{-strong convexity}) \\
&= -\eta\, \delta f_t(\mathbf{w}_t) \cdot (\mathbf{w}_t - \mathbf{w}_{t+1}) - \frac{\alpha}{2}\|\mathbf{w}_t - \mathbf{w}_{t+1}\|^2 && (\text{def. of } \mathbf{v}_{t+1}) \\
&\leq \eta G_* \|\mathbf{w}_t - \mathbf{w}_{t+1}\| - \frac{\alpha}{2}\|\mathbf{w}_t - \mathbf{w}_{t+1}\|^2 && (G_*\text{-Lipschitzness}) \\
&\leq \frac{(\eta G_*)^2}{2\alpha}. && (\text{max. of 2nd deg. eq.})
\end{aligned}
$$

- **Theorem**: assume additionally that f_ts are σ-strongly convex with respect to Φ. Then, the regret of Mirror Descent with parameter $\eta_{t+1} = \frac{1}{\sigma t}$ can be bounded as follows:

$$R_T(\mathrm{MD}) \leq \frac{G_*^2}{2\sigma\alpha}(1 + \log T).$$

Equivalent Description of OMD

MIRROR-DESCENT(Φ)

$\mathbf{w}_1 \leftarrow \mathrm{argmin}_{\mathbf{w}\in K\cap C}\, \Phi(\mathbf{w})$
for $t \leftarrow 1$ **to** $(T-1)$ **do**
  $\mathbf{w}_{t+1} \leftarrow \mathrm{argmin}_{\mathbf{w}\in K\cap C}\, \underbrace{\delta f_t(\mathbf{w}_t)\cdot \mathbf{w}}_{\text{linearization of } f_t} + \underbrace{\frac{1}{\eta} B(\mathbf{w} \parallel \mathbf{w}_t)}_{\text{regularization}}$

- Proof:

$$
\begin{aligned}
\mathbf{w}_{t+1} &= \operatorname*{argmin}_{\mathbf{w}\in K\cap C}\ B(\mathbf{w} \parallel \mathbf{v}_{t+1}) \\
&= \operatorname*{argmin}_{\mathbf{w}\in K\cap C}\ \Phi(\mathbf{w}) - \nabla\Phi(\mathbf{v}_{t+1})\cdot \mathbf{w} && \text{(def. of Breg. div.)} \\
&= \operatorname*{argmin}_{\mathbf{w}\in K\cap C}\ \Phi(\mathbf{w}) - \big(\nabla\Phi(\mathbf{w}_t) - \eta\,\delta f_t(\mathbf{w}_t)\big)\cdot \mathbf{w} && \text{(def. of } \mathbf{v}_{t+1}\text{)} \\
&= \operatorname*{argmin}_{\mathbf{w}\in K\cap C}\ \eta\,\delta f_t(\mathbf{w}_t)\cdot \mathbf{w} + B(\mathbf{w} \parallel \mathbf{w}_t). && \text{(def. of Breg. div.)}
\end{aligned}
$$

Dual Averaging (DA) [Louditski and Nesterov, 2010]

DUAL-AVERAGING(Φ)

$\mathbf{v}_1 \leftarrow 0$
$\mathbf{w}_1 \leftarrow \operatorname{argmin}_{\mathbf{w} \in K \cap C} B(\mathbf{w} \parallel \mathbf{v}_1)$
for $t \leftarrow 1$ **to** T **do**
$\qquad \mathbf{v}_{t+1} \leftarrow [\nabla \Phi]^{-1}\big(\nabla \Phi(\mathbf{v}_t) - \eta\, \delta f_t(\mathbf{w}_t)\big)$
$\qquad \mathbf{w}_{t+1} \leftarrow \operatorname{argmin}_{\mathbf{w} \in K \cap C} B(\mathbf{w} \parallel \mathbf{v}_{t+1})$

Dual Averaging (DA) [Louditski and Nesterov, 2010]

DUAL-AVERAGING(Φ)

$\mathbf{v}_1 \leftarrow 0$
$\mathbf{w}_1 \leftarrow \operatorname{argmin}_{\mathbf{w} \in K \cap C} B(\mathbf{w} \parallel \mathbf{v}_1)$
for $t \leftarrow 1$ **to** T **do**
  $\mathbf{v}_{t+1} \leftarrow [\nabla \Phi]^{-1}(\nabla \Phi(\mathbf{v}_t) - \eta\, \delta f_t(\mathbf{w}_t))$
  $\mathbf{w}_{t+1} \leftarrow \operatorname{argmin}_{\mathbf{w} \in K \cap C} B(\mathbf{w} \parallel \mathbf{v}_{t+1})$

$\nabla \Phi(v_t) = \sum \cdots$

A simple modification makes a big difference

MIRROR-DESCENT(Φ)

$\mathbf{w}_1 \leftarrow \operatorname{argmin}_{\mathbf{w} \in K \cap C} \Phi(\mathbf{w})$
for $t \leftarrow 1$ **to** T **do**
  $\mathbf{v}_{t+1} \leftarrow [\nabla \Phi]^{-1}(\nabla \Phi(\mathbf{w}_t) - \eta\, \delta f_t(\mathbf{w}_t))$
  $\mathbf{w}_{t+1} \leftarrow \operatorname{argmin}_{\mathbf{w} \in K \cap C} B(\mathbf{w} \parallel \mathbf{v}_{t+1})$

Dual Averaging (DA) [Louditski and Nesterov, 2010]

Dual-Averaging(Φ)

$\mathbf{v}_1 \leftarrow 0$

$\mathbf{w}_1 \leftarrow \operatorname{argmin}_{\mathbf{w}\in K\cap C} B(\mathbf{w} \parallel \mathbf{v}_1)$

for $t \leftarrow 1$ **to** T **do**

$\quad \mathbf{v}_{t+1} \leftarrow [\nabla\Phi]^{-1}\left(\nabla\Phi(\mathbf{v}_t) - \eta\, \delta f_t(\mathbf{w}_t)\right)$

$\quad \mathbf{w}_{t+1} \leftarrow \operatorname{argmin}_{\mathbf{w}\in K\cap C} B(\mathbf{w} \parallel \mathbf{v}_{t+1})$

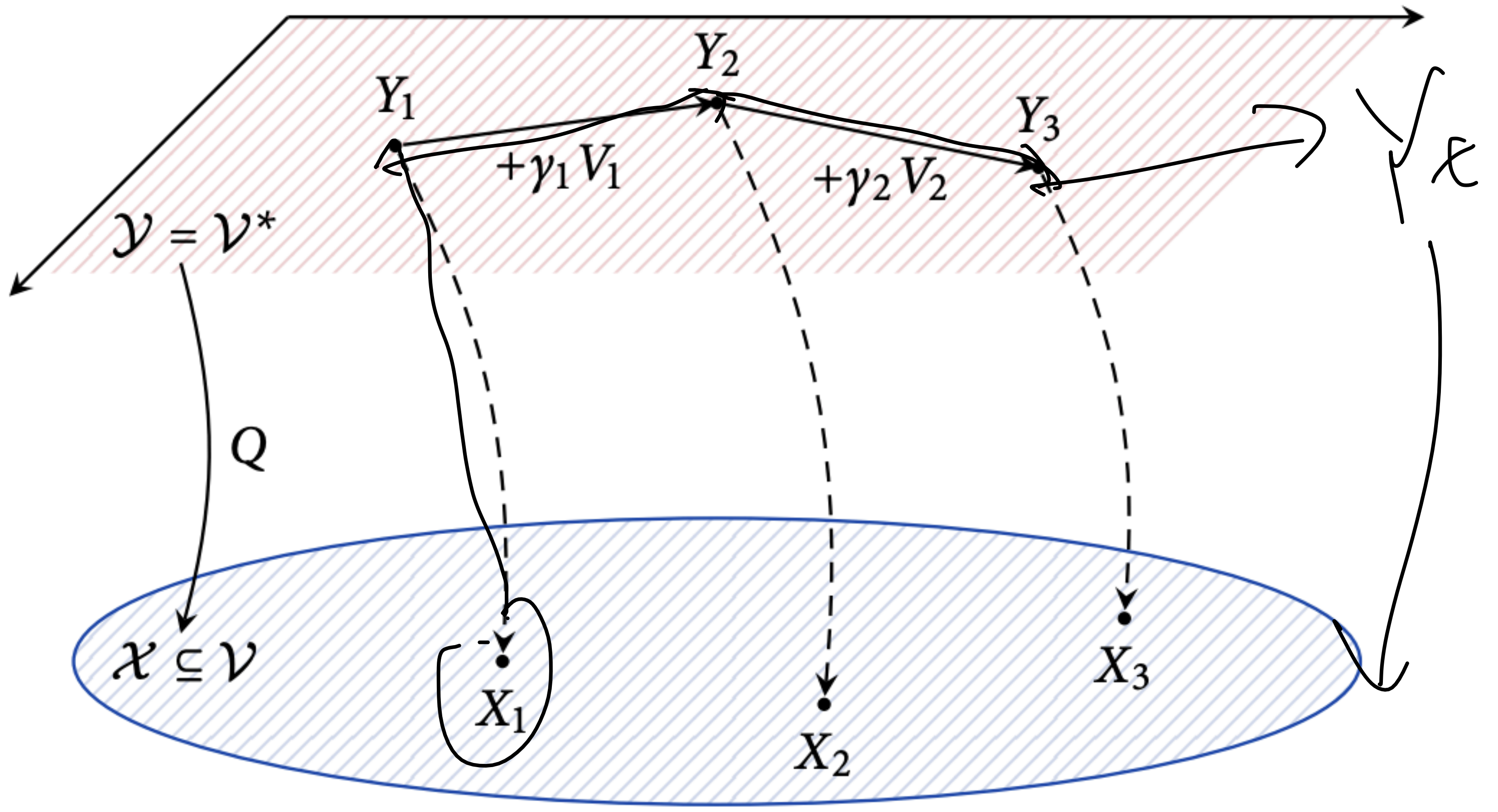


Figure 2.3: Schematic representation of dual averaging.

Mirror-Descent(Φ)

$\mathbf{w}_1 \leftarrow \operatorname{argmin}_{\mathbf{w}\in K\cap C} \Phi(\mathbf{w})$

for $t \leftarrow 1$ **to** T **do**

$\quad \mathbf{v}_{t+1} \leftarrow [\nabla\Phi]^{-1}\left(\nabla\Phi(\mathbf{w}_t) - \eta\, \delta f_t(\mathbf{w}_t)\right)$

$\quad \mathbf{w}_{t+1} \leftarrow \operatorname{argmin}_{\mathbf{w}\in K\cap C} B(\mathbf{w} \parallel \mathbf{v}_{t+1})$

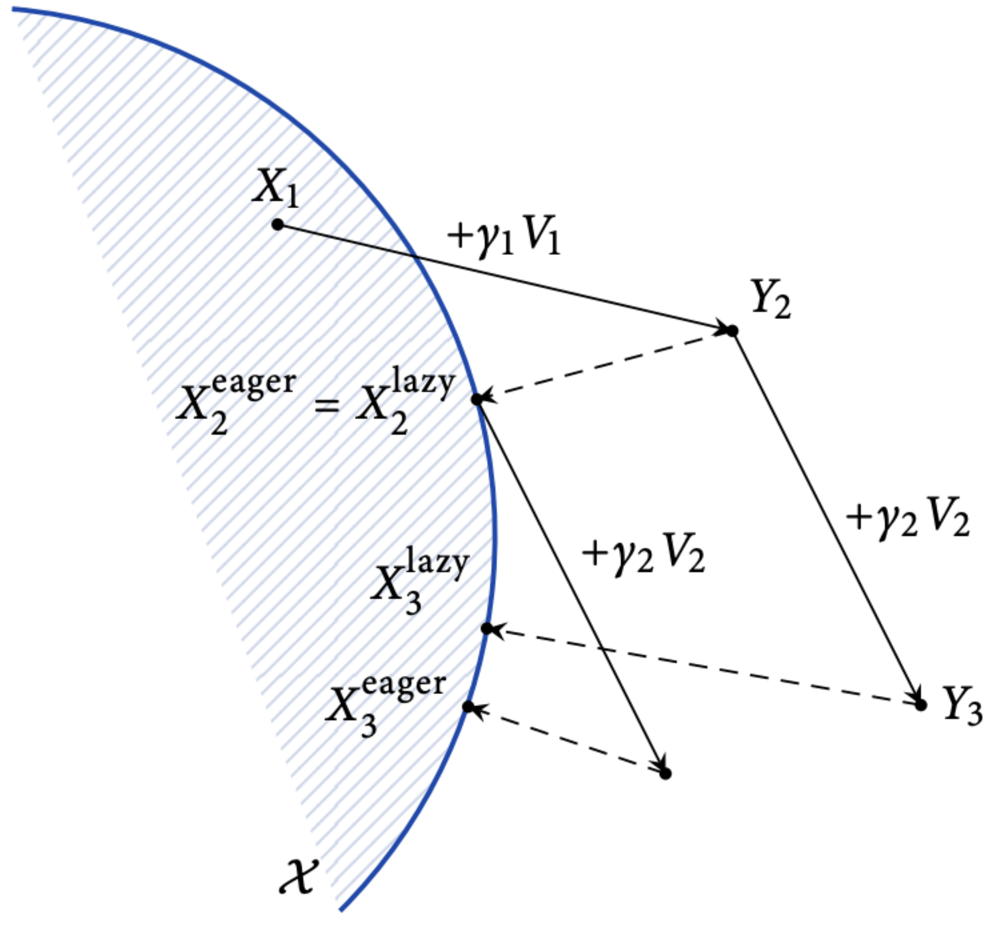


Figure 2.4: Lazy vs. eager gradient descent.

Equivalent Description of DA

- Equivalent form:

$$\begin{aligned}
\mathbf{w}_{t+1} &= \operatorname*{argmin}_{\mathbf{w}\in K\cap C} \; B(\mathbf{w} \,\|\, \mathbf{v}_{t+1}) \\
&= \operatorname*{argmin}_{\mathbf{w}\in K\cap C} \; \Phi(\mathbf{w}) - \nabla\Phi(\mathbf{v}_{t+1})\cdot\mathbf{w} && \text{(def. of Breg. div.)} \\
&= \operatorname*{argmin}_{\mathbf{w}\in K\cap C} \; \Phi(\mathbf{w}) - \big(\nabla\Phi(\mathbf{v}_t) - \eta\,\delta f_t(\mathbf{w}_t)\big)\cdot\mathbf{w} && \text{(def. of } \mathbf{v}_{t+1}\text{)} \\
&= \operatorname*{argmin}_{\mathbf{w}\in K\cap C} \; \eta\sum_{s=1}^{t}\delta f_t(\mathbf{w}_s) + \Phi(\mathbf{w}). && \text{(recurrence)}
\end{aligned}$$

- In particular, for linear losses, $f_t(\mathbf{w}) = \mathbf{a}_t \cdot \mathbf{w}$, Dual Averaging coincides with regularized FL:

$$\mathbf{w}_{t+1} = \operatorname*{argmin}_{\mathbf{w}\in K\cap C} \; \sum_{s=1}^{t}\mathbf{a}_s\cdot\mathbf{w} + \frac{1}{\eta}\Phi(\mathbf{w}).$$

Comparison between OMD and DA

OMD

$$\begin{aligned}
\mathbf{w}_{t+1} &= \operatorname*{argmin}_{\mathbf{w}\in K\cap C} \; B(\mathbf{w} \,\|\, \mathbf{v}_{t+1}) \\
&= \operatorname*{argmin}_{\mathbf{w}\in K\cap C} \; \Phi(\mathbf{w}) - \nabla\Phi(\mathbf{v}_{t+1}) \cdot \mathbf{w} \\
&= \operatorname*{argmin}_{\mathbf{w}\in K\cap C} \; \Phi(\mathbf{w}) - \big(\nabla\Phi(\mathbf{w}_t) - \eta\,\delta f_t(\mathbf{w}_t)\big) \cdot \mathbf{w} \\
&= \operatorname*{argmin}_{\mathbf{w}\in K\cap C} \; \eta\,\delta f_t(\mathbf{w}_t) \cdot \mathbf{w} + B(\mathbf{w} \,\|\, \mathbf{w}_t).
\end{aligned}$$

w

DA

$$\begin{aligned}
\mathbf{w}_{t+1} &= \operatorname*{argmin}_{\mathbf{w}\in K\cap C} \; B(\mathbf{w} \,\|\, \mathbf{v}_{t+1}) \\
&= \operatorname*{argmin}_{\mathbf{w}\in K\cap C} \; \Phi(\mathbf{w}) - \nabla\Phi(\mathbf{v}_{t+1}) \cdot \mathbf{w} \\
&= \operatorname*{argmin}_{\mathbf{w}\in K\cap C} \; \Phi(\mathbf{w}) - \big(\nabla\Phi(\mathbf{v}_t) - \eta\,\delta f_t(\mathbf{w}_t)\big) \cdot \mathbf{w} \\
&= \operatorname*{argmin}_{\mathbf{w}\in K\cap C} \; \eta \sum_{s=1}^{t} \delta f_t(\mathbf{w}_s) + \Phi(\mathbf{w}).
\end{aligned}$$

Regret of DA

- Theorem: under the same assumptions as for MD, the following holds for the regret of Dual Averaging,

$$R_T(\mathrm{DA}) \leq \frac{\Phi(\mathbf{w}^*) - \Phi(\mathbf{w}_1)}{\eta} + \frac{2\eta G_*^2 T}{\alpha}.$$

Choosing η to minimize the bound gives

$$R_T(\mathrm{DA}) \leq 2 D_\Phi G_* \sqrt{\frac{2T}{\alpha}},$$

with $\Phi(\mathbf{w}^*) - \Phi(\mathbf{w}_1) \leq D_\Phi^2$.

Summary

- Online Convex Optimization
  1. Full information feedback (FTRL)
  2. First-order feedback
     A. OPSD/EG as incarnations of OMD
     B. From OMD to DA
  3. Zeroth-order feedback
     (adversarial bandit problem)
     C. For pseudo-regret, EXP3 as a modification of EG.
     D. For true regret, EXP3-P, EXP3-IX.