Polynomial Weights Algorithm

$$w_i^1 \leftarrow 1 \quad \forall \text{ experts } i \in \{1, \ldots, N\}$$

for each round $t$ in $1 \ldots T$:

$$w^t \leftarrow \sum_{i=1}^N w_i^t$$

choose expert $i$ w/prob $\frac{w_i^t}{W^t}$

$$w_i^{t+1} \leftarrow w_i^t \cdot (1 - \varepsilon l_i^t) \quad \forall i \in \{1, \ldots, N\}$$

end for

Bound Algorithm Loss

$$f(L_{opt}^T) \leq W^{T+1} \leq g(L_T^T)$$

Total loss of the best expert

Total weight of all experts at the end

Total loss of the algorithm
Last time:

\[ \ln(W^{T+1}) \leq \ln(N) - \varepsilon \mathbb{E}(L^T) \]

Still need:

\[ W^{T+1} \geq ? \]

\[ W^{T+1} \geq w_i^{T+1} \quad \forall \text{ experts } i \]

\[ w_i^{T+1} = \prod_{t=1}^{T} (1 - \varepsilon l_i^t) \]

\[ \ln(W^{T+1}) \geq \sum_{t=1}^{T} \ln(1 - \varepsilon l_i^t) \]
using \( \ln (1-x) \geq -x - x^2 \)
for \( 0 \leq x \leq 0.5 \)

\[
\ln (W^T) \geq -\sum_{t=1}^{T} \varepsilon L_i^t - \sum_{t=1}^{T} (\varepsilon L_i^t)^2 \\
\geq -\varepsilon L_i^T - \varepsilon^2 T
\]
for $0 \leq \kappa \leq 0.5$

$$\ln(1-\kappa) \geq -\kappa - \kappa^2$$
Combine:

\[-3L_i^T - \varepsilon^2 T \leq \ln(N) - \varepsilon \mathbb{E}(L^T)\]

\[
\min_i L_i^T \quad \text{(best expert i)}
\]

\[
\varepsilon \mathbb{E}(L^T) \leq 3L_{opt}^T + \varepsilon^2 T + \ln(N)
\]

\[
\mathbb{E}(L^T) \leq L_{opt}^T + \varepsilon T + \frac{\ln(N)}{\varepsilon}
\]

Let's choose \( \varepsilon = \sqrt{\frac{\ln(N)}{T}} \)

\[
\mathbb{E}(L^T) \leq L_{opt}^T + 2\sqrt{T \ln(N)}
\]
\[
\frac{E(L^T)}{T} \leq \frac{L^T_{\text{opt}}}{T} + 2 \sqrt{\frac{\ln(N)}{T}}
\]

PW vs. Safety value?
Recall: A minimax strategy is a best response to the worst-case opponent profile:

$$\arg\min_{\overrightarrow{\sigma_i}} u_i(BR_i(\overrightarrow{\sigma_{-i}}), \overrightarrow{\sigma_i})$$

The player's safety value is the corresponding utility:

$$\min_{\overrightarrow{\sigma_i}} u_i(BR_i(\overrightarrow{\sigma_{-i}}), \overrightarrow{\sigma_i})$$
Minimax Theorem

In any 2-player zero-sum game, player 1's safety value $SV_1$ is minus player 2's safety value $SV_2$

$$SV_1 = -SV_2$$

Proof:
Consider the strategies $\sigma_1, \sigma_2$ that lead to P1's safety value

$$u_1 (\sigma_1, \sigma_2) = SV_1$$

$\sigma_1$ is a best response to $\sigma_2$

$$u_2 (\sigma_2, \sigma_1) = -SV_1$$
$\sigma_2$ is a BR to $\sigma_1$.

If $p_2$ pick $\sigma_2$, then $-SV_1$ is the worst-case, so

$$SV_2 \geq -SV_1.$$ 

Now we need to show $SV_2 \neq -SV_1$.
Suppose it were ($SV_2 > -SV_1$)
Then $\exists \epsilon > 0 \text{ s.t. } SV_2 = -SV_1 + \epsilon$

Let $p_1$ follow $PW + p_2$ best-respond...

What can we say about $p_1$'s payoff? $p_2$'s?