$$E_\xi \Gamma(f_{t+1}) - \Gamma(f_t) \le -\frac{\alpha\mu}{2}\|\nabla\Gamma(f_t)\|^2 + \frac{\alpha^2 MB}{2}$$

Theorem 10.9 suppose Γ is M smooth, $0 < \alpha \le \frac{\mu}{MB\Gamma}$

+ Assumption 10.1

then $$E\,\frac{1}{T}\sum_{t=1}^{T}\|\nabla\Gamma(f_t)\|^2 \le \frac{\alpha MB}{\mu} + \frac{2(\Gamma(f_1) - \Gamma(f^*))}{T\mu\alpha}$$

Proof

$$\overbrace{E_\xi \Gamma(f_{t+1}) - \Gamma(f_t)}^{E} \le -\frac{\alpha\mu}{2}\|\nabla\Gamma(f_t)\|^2 + \frac{\alpha^2 MB}{2}$$

$$\|\nabla\Gamma(f_t)\|^2 \le \frac{2}{\alpha\mu}\left[E\Gamma(f_t) - E\Gamma(f_{t+1})\right] + \frac{\alpha MB}{\mu}$$

$$\frac{1}{T}\sum_{t=1}^{T}\|\nabla\Gamma(f_t)\|^2 \le \frac{2}{\alpha\mu T}\left(E\Gamma(f_1) - \underbrace{\cancel{E\Gamma(f_{T+1})}}_{\Gamma^*}\right) + \frac{\alpha MB}{\mu}$$

$\square$

Zinkevich - Chapt. 11, Sect. 1
Adversarial approach to online convex programming

<u>Setting</u> Let

$\mathcal{F}$ - closed convex subset of a Hilbert space $\mathcal{H}$

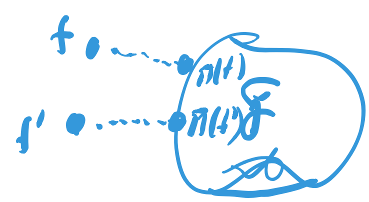
$D = \max\{\|f-f'\| : f, f' \in \mathcal{F}\} < +\infty$

$\ell : \mathcal{F} \times Z \to \mathbb{R}$ $\quad \ell(f_t, z_t)$ = loss at time t.

Z - an arbitrary set, set of possible z_t's

player can compute $\ell(f, z_t)$ and $\nabla \ell(f, z_t)$ for given f.

player can use projection $\Pi : \mathcal{H} \to \mathcal{F}$
(<u>Note</u> $\|\Pi(f) - \Pi(f')\| \le \|f - f'\|$)



Online game:
- At each time t, player selects $f_t \in \mathcal{F}$
- Then adversary selects $z_t \in Z$
- Player sees z_t, incurs loss $\ell(f_t, z_t)$.

$$J_T((f_t)) = \sum_{t=1}^{T} \ell(f_t, z_t) \qquad L_T((f_t)) = \frac{1}{T} J_T((f_t))$$

<u>regret</u>

$$R_T((f_t)) = J_T((f_t)) - \inf_{f^*} J_T(f^*)$$

Consider projected gradient descent algorithm

f_1 - given $\qquad f_{t+1} = \Pi\left(f_t - \alpha_t \nabla \ell(f_t, z_t)\right)$

<u>Thm 11.1</u> Suppose $\ell(\cdot, z)$ is convex and L-Lipschitz continuous for each z. Use projected gradient

(a) If $\alpha_t = \frac{c}{\sqrt{t}}$ for $t \geq 1$ then

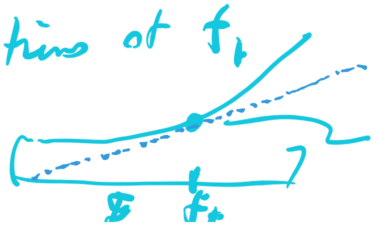
$$R_T((f_t)) \leq \frac{D^2\sqrt{T}}{2c} + \left(\sqrt{T} - \frac{1}{2}\right)L^2 c$$

(b) If in addition, $\ell(\cdot, z)$ is m-strongly convex for some $m > 0$, and $\alpha_t = \frac{1}{mt}$ then

$$R_T((f_t)) \leq \frac{L^2(1 + \log T)}{2m}$$

<u>Observation</u>: For purpose of proof of part (a) we could assume that the functions $\ell(\cdot, z)$ are all linear functions of f_t

Why?



$\ell(f_t, z_t)$

Proof Let $f^b_{t+1} = f_t - \alpha_t \nabla l_t(f_t)$ $\quad (\nabla l(f_t, z_t))$

$$f_{t+1} = \Pi(f^b_{t+1})$$

Note $f^b_{t+1} - f^* = \underline{f_t - f^*} - \underline{\alpha_t \nabla l_t(f_t)}$

$$\|f_{t+1} - f^*\|^2 \leq \|f^b_{t+1} - f^*\|^2 \leq \|f_t - f^*\|^2 - 2\alpha_t \langle f_t - f^*, \nabla l_t(f_t)\rangle + \alpha_t^2 L^2$$

so

$$2\langle f_t - f^*, \nabla l_t(f_t)\rangle \leq \frac{\|f_t - f^*\|^2 - \|f_{t+1} - f^*\|^2}{\alpha_t} + \alpha_t L^2$$

seeds of telescoping sum

To be continued.