$F: \mathbb{R}^d \to \mathbb{R}^d$ $\qquad$ Z - convex set

Find $z^* \in Z$ $\qquad$ $\langle F(z^*); z - z^* \rangle \geq 0 \quad \forall z \in Z$

Variational iniqulity problem

$Z = \mathbb{R}^d$ $\qquad$ $\langle F(z^*); z - z^* \rangle \geq 0 \quad \forall z \in \mathbb{R}^d$

$F(z^*) \neq \vec{0}$ $\qquad$ $z = z^* - F(z^*) \in \mathbb{R}^d$

$$\langle F(z^*); -F(z^*) \rangle = -\|F(z^*)\|^2 < 0$$

$$F(z^*) = 0$$

1) $\nabla f(z^*) = 0$

2) $F(z) = \begin{pmatrix} \nabla_x f(x,y) \\ -\nabla_y f(x,y) \end{pmatrix}$ $\qquad$ $z = \begin{pmatrix} x \\ y \end{pmatrix}$

$f(x,y)$ - convex for x (for any fixed y)

- concave for y (for any fixed x)

VI for this F

$$\left\langle \begin{pmatrix} \nabla_x f(x^*, y^*) \\ -\nabla_y f(x^*, y^*) \end{pmatrix}; \begin{pmatrix} x - x^* \\ y - y^* \end{pmatrix} \right\rangle \geq 0$$

$$= \underbrace{\langle \nabla_x f(x^*; y^*); x - x^* \rangle}_{} + \underbrace{\langle -\nabla_y f(x^*; y^*); y - y^* \rangle}_{} \geq 0$$

$$\underline{f(x, y^*) - f(x^*; y^*) + f(x^*; y^*) - f(x^*; y) \geq 0}$$

a) $x = x^* \rightarrow \forall y \quad f(x^*; y^*) - f(x^*; y) \geq 0$

d) $y = y^* \rightarrow \forall x \quad f(x; y^*) - f(x^*; y^*) \geq 0$

saddle point

$$\boxed{\min_x \max_y f(x, y)}$$

Example 1 games

$$\min_{x \in \Delta} \max_{y \in \Delta} (x^T A y)$$

A_{ij} – payments (if y choose action j, x choose action i)

$$\Delta = \{ x \in \mathbb{R}^d \mid x_i \geq 0 \quad \sum x_i = 1 \}$$

Example 2 Adversarial noise

$$\min_{x \in \mathbb{R}^d} \max_{\|y_i\| \leq \delta} \frac{1}{n} \sum l(g(x, a_i + y_i); b_i)$$

Example 3 Lagrange mult.

$$\min_{x \in \mathbb{R}^d} f(x) \quad \text{s.t.} \quad g(x) = 0, \; h(x) \leq 0 \quad \Rightarrow \quad \min_x \max_{y, \lambda} [f(x) + y g(x) + \lambda h(x)]$$

1) $\|F(z_1) - F(z_2)\| \leq L \|z_1 - z_2\|$

$\|\nabla f(z_1) - \nabla f(z_2)\| \leq L \|z_1 - z_2\|$ L-smoothness

2) $\langle F(z_1) - F(z_2); z_1 - z_2 \rangle \geq \mu \|z_1 - z_2\|^2$

μ-strong convexity of $f(z)$

μ-strong convexity - strong concavity of $f(x,y)$

For min:

$z^{k+1} = z^k - \gamma \nabla f(z^k) \Rightarrow z^{k+1} = z^k - \gamma F(z^k)$

$Z = \mathbb{R}^d \qquad F(z^*) = 0$

$$\|z^{k+1} - z^*\|^2 = \|z^k - \gamma F(z^k) - z^*\|^2$$
$$= \|z^k - z^*\|^2 - 2\gamma \langle F(z^k); z^k - z^* \rangle + \gamma^2 \|F(z^k)\|^2$$

$F(z^*)=0$

$$= \|z^k - z^*\|^2 - 2\gamma \langle F(z^k) - F(z^*); z^k - z^* \rangle + \gamma^2 \|F(z^k) - F(z^*)\|^2$$

Assumptions

$$\leq \|z^k - z^*\|^2 - 2\gamma\mu \|z^k - z^*\|^2 + \gamma^2 L^2 \|z^k - z^*\|^2$$

$$= \underbrace{(1 - 2\gamma\mu + \gamma^2 L^2)}_{< 1} \|z^k - z^*\|^2$$

$(1 - 2\gamma\mu + \gamma^2 L^2)' = 0 \qquad -2\mu + 2\gamma L^2 = 0$

$\gamma = \mu / L^2$

$$= \left(1 - \frac{2\mu^2}{L^2} + \frac{\mu^2}{L^2}\right)\|z^k - z^*\|^2$$

$$= \left(1 - \frac{\mu^2}{L^2}\right)\|z^k - z^*\|^2$$

$$O\left(\frac{L^2}{\mu^2}\log\frac{1}{\varepsilon}\right)$$ good?

GD $O\left(\frac{L}{\mu}\log\frac{1}{\varepsilon}\right)$

AccGD $O\left(\sqrt{\frac{L}{\mu}}\log\frac{1}{\varepsilon}\right)$

---

why method is not good

$$\min_{x\in\mathbb{R}}\max_{y\in\mathbb{R}} xy \qquad x^* = 0,\quad y^* = 0$$

$$\begin{pmatrix} \nabla_x f(x,y) = y \\ -\nabla_y f(x,y) = -x \end{pmatrix} = 0 \text{ in } (x^*; y^*)$$

$$x^{k+1} = x^k - \gamma \nabla_x f(x^k, y^k)$$

$$y^{k+1} = y^k + \gamma \nabla_y f(x^k, y^k)$$

$$\|z^{k+1} - z^*\|^2 = (x^{k+1} - \underbrace{x^*}_{=0})^2 + (y^{k+1} - \underbrace{y^*}_{=0})^2$$

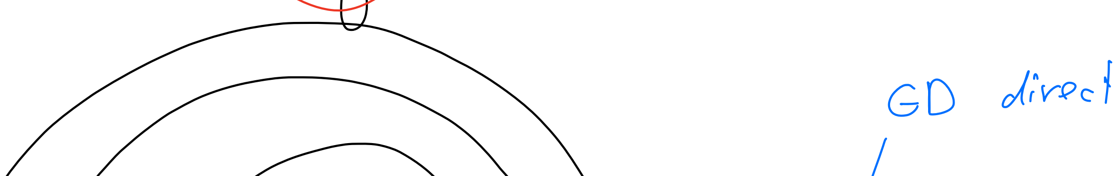
$$= (x^k - \gamma y^k)^2 + (y^k + \gamma x^k)^2$$

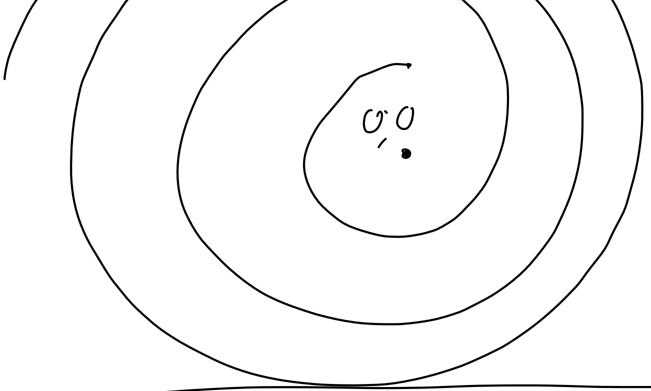
$$= (x^k)^2 + (y^k)^2 + \gamma^2 (y^k)^2 + \gamma^2 (x^k)^2 - \cancel{2\gamma x^k y^k} + \cancel{2\gamma x^k y^k}$$

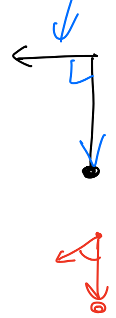
$$= (1+\gamma^2)(x^k)^2 + (1+\gamma^2)(y^k)^2$$

$$= (1+\gamma^2)(x^k - x^*)^2 + (1+\gamma)^2(y^k - y^*)^2$$

$$= (1+\gamma^2)\|z^k - z^*\|^2$$







choose at home

Extra Gradient

$$z^{k+1/2} = z^k - \gamma F(z^k)$$
$$z^{k+1} = z^k - \gamma F(z^{k+1/2})$$

Proof:

Lemma $z^+ = z - y$ $\forall u \in \mathbb{R}^d$

$$\|z^+ - u\|^2 = \|z^+ - z + z - u\|^2$$
$$= \|z-u\|^2 + \|z^+ - z\|^2 + 2\langle z^+ - z; z - u\rangle$$
$$= \|z-u\|^2 + \|z^+ - z\|^2 + 2\langle z^+ - z; z - z^+ + z^+ - u\rangle$$
$$= \|z-u\|^2 + \|z^+ - z\|^2 - 2\|z^+ - z\|^2 + 2\langle z^+ - z; z^+ - u\rangle$$
$$= \|z-u\|^2 - \|z^+ - z\|^2 - 2\langle y; z^+ - u\rangle \quad (+)$$

1) $z^+ = z^{k+1}$ $y = \gamma F(z^{k+1/2})$ $u = z^*$ $z = z^k$

$$+\ \|z^{k+1} - z^*\|^2 = \|z^k - z^*\|^2 - \cancel{\|z^{k+1} - z^k\|^2} - 2\gamma\langle F(z^{k+1/2}); z^{k+1} - z^*\rangle$$

2) $z^+ = z^{k+1/2}$ $y = \gamma F(z^k)$ $u = z^{k+1}$ $z = z^k$

$$+\ \|z^{k+1/2} - z^{k+1}\|^2 = \cancel{\|z^k - z^{k+1}\|^2} - \|z^{k+1/2} - z^k\|^2 - 2\gamma\langle F(z^k); z^{k+1/2} - z^{k+1}\rangle$$

$$\|z^{k+1} - z^*\|^2 + \|z^{k+1/2} - z^{k+1}\|^2 = \|z^k - z^*\|^2 - 2\gamma\langle F(z^{k+1/2}); z^{k+1} - z^*\rangle - \|z^{k+1/2} - z^k\|^2 - 2\gamma\langle F(z^k); z^{k+1/2} - z^{k+1}\rangle$$

$$-2\gamma\langle F(z^{k+1/2}), z^{k+1/2}-z^*\rangle - 2\gamma\langle F(z^k)-F(z^{k+1/2}), z^{k+1/2}-z^{k+1}\rangle$$

$-F(z^*) = 0$

CS

$$\le \gamma^2\|F(z^k)-F(z^{k+1/2})\|^2 + \|z^{k+1/2}-z^{k+1}\|^2$$

L-Lipsch

$$\le -2\mu\gamma\|z^{k+1/2}-z^*\|^2 + \gamma^2L^2\|z^k-z^{k+1/2}\|^2 + \|z^{k+1/2}-z^{k+1}\|^2$$

$$\|z^{k+1}-z^*\|^2 \le \|z^k-z^*\|^2 - \|z^{k+1/2}-z^k\|^2 + \cancel{\|z^{k+1/2}-z^{k+1}\|^2}$$
$$\cancel{\|z^{k+1/2}-z^{k+1}\|^2} \quad - 2\mu\gamma\|z^{k+1/2}-z^*\|^2 + \gamma^2L^2\|z^k-z^{k+1/2}\|^2$$

$$\|z^{k+1}-z^*\|^2 \le \|z^k-z^*\|^2 - 2\mu\gamma\|z^{k+1/2}-z^*\|^2 - (1-\gamma^2L^2)\|z^k-z^{k+1/2}\|^2$$

$$\|z^{k+1/2}-z^*\|^2 \longrightarrow \|z^k-z^*\|^2$$

CS

$$\|z^k-z^*\|^2 \le 2\|z^k-z^{k+1/2}\|^2 + 2\|z^{k+1/2}-z^*\|^2$$

$-z^{k+1/2}+z^{k+1/2}$

$$-2\|z^{k+1/2}-z^*\|^2 \le -\|z^k-z^*\|^2 + 2\|z^k-z^{k+1/2}\|^2$$

$$\|z^{k+1}-z^*\|^2 \le (1-\mu\gamma)\|z^k-z^*\|^2$$
$$-(1-\gamma^2L^2-2\mu\gamma)\|z^k-z^{k+1/2}\|^2$$

Linear converg

$$\gamma = \min\left(\frac{1}{\sqrt{2}L}; \frac{1}{2\mu}\right)$$

$$\gamma^2L^2 \le \frac{1}{2} \qquad 2\mu\gamma \le \frac{1}{2}$$

$$\boxed{\|z^{k+1} - z^*\|^2 \le (1-\mu\gamma)\|z^k - z^*\|^2}$$

$\gamma = \frac{1}{L};\ \frac{1}{\mu}$

$O\left(\left(1+\frac{L}{\mu}\right)\log\frac{1}{\varepsilon}\right)$

(+) GD $\frac{L}{\mu}\log\frac{1}{\varepsilon}$

(?) AccGD $\sqrt{\frac{L}{\mu}}\log\frac{1}{\varepsilon}$

unimprovable

no acceleration

$z^{k+1/2} = z^k - \gamma F(z^{k-1/2})$ ← computed on prev iter

$z^{k+1} = z^k - \gamma F(z^{k+1/2})$

Optimistic EG

only 1 call of F (EG calls 2 times F)