Projected Gradient Descent

- Projection onto closed convex sets
- Projected Gradient descent.

Last time

Projection.

The orthogonal projection $\text{proj}_C : \mathbb{R}^n \to C$

$$\text{proj}_C(x) = \underset{z \in C}{\arg\min}\ p(z) = \arg\min \left\{ \|z - x\|_2^2 \;/\; z \in C \right\}.$$

- Returns a point in the set C closest to $x \in \mathbb{R}^n$.

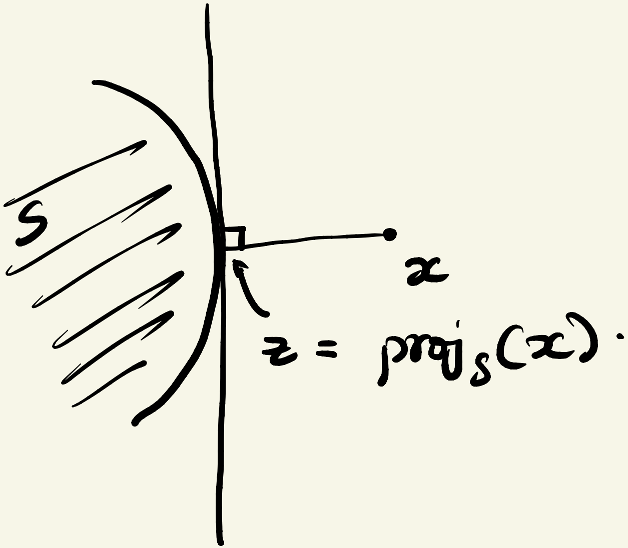
- If $x \in C$, $\text{proj}_C(x) = x$.

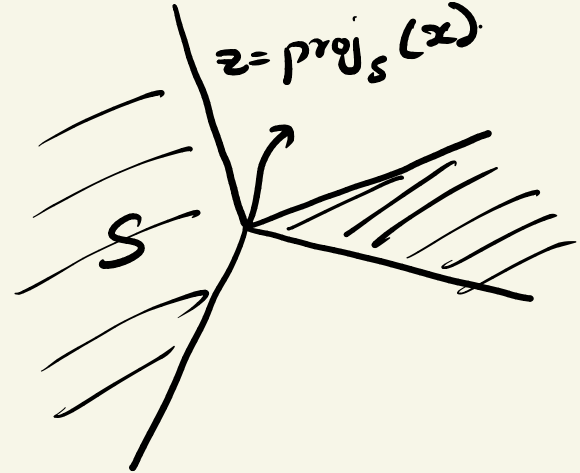
- Optimality conditions for projection: Since $\nabla p(x) = z - x$

$$z = \text{proj}_C(x) \iff (z-x)^T(y-z) \geq 0 \quad \forall y \in C.$$
$$\iff x - z \in N_C(z).$$

$$N_C(z) = \left\{ g \;/\; g^T(z-y) \leq 0, \ \forall y \in C \right\}.$$

Geometry of projection.



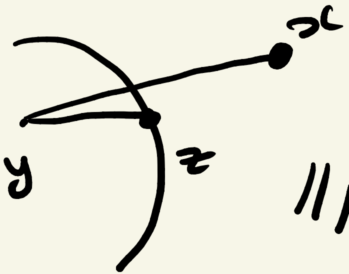


$$N_S(z) = \{\lambda \cdot (x - z) \mid \lambda \geq 0\}.$$

Thm

Projection onto a closed convex set is unique.

Contractive property



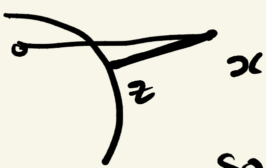
$$\|\mathrm{proj}_S(x) - y\|_2 \le \|x - y\|_2 \quad \forall\, y \in S.$$

<u>Proof</u>: Take $z = \mathrm{proj}_S(x)$. By optimality condition for projection, we have

$$0 \le (z-x)^T(y-z) = (z-x)^T(y-z+x-x)$$

$$= -\|z-x\|_2^2 + (z-x)^T(y-x)$$

$$\le -\|z-x\|_2^2 + \|z-x\|_2\,\|y-x\|_2 \qquad \left[\|z-x\|_2 \le \|y-x\|_2\right]$$

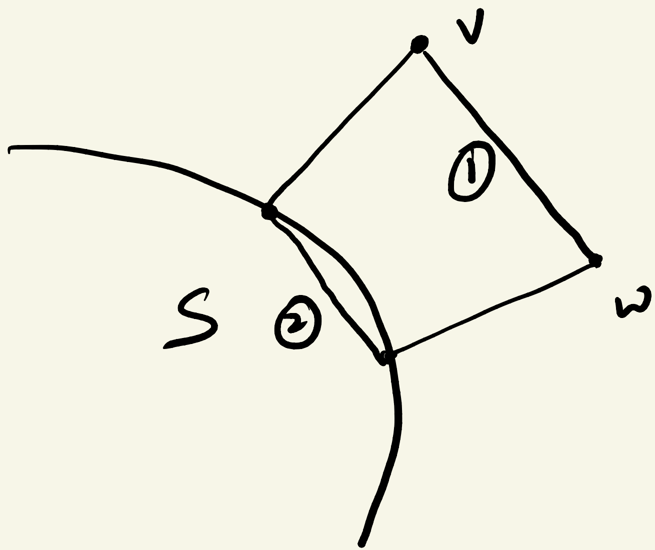


$$\le -\|z-x\|_2^2 + \|y-x\|_2^2$$

so, $\|z-x\|_2^2 \le \|y-x\|_2^2 \quad \forall\, y \in S.$

Thm: Let C be a closed convex set. Then for any $v, w \in \mathbb{R}^n$,

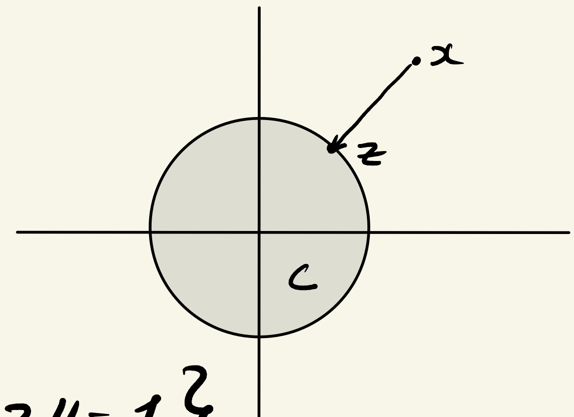
$$\|proj_C(v) - proj_C(w)\|_2 \leq \|v - w\|_2$$



① > ②

Examples

$$z = \mathrm{proj}_C(x) \quad , \qquad C = \{ z \mid \|z\|_2 \le 1 \}$$



- If $x \in C$, $\mathrm{proj}_C(x) = x$
- If $\|x\|_2 > 1$, $z = \dfrac{1}{\|x\|_2} x$

proof sketch: If $\|x\|_2 > 1$,

$$\begin{aligned} z &= \mathrm{argmin} \{ \|z - x\|_2^2 \mid \|z\| = 1 \} \\ &= \mathrm{argmin} \{ -2x^T z + \|z\|_2^2 + \|x\|_2^2 \mid \|z\| = 1 \}. \\ &= \mathrm{argmin} \{ -2z^T x \mid \|z\|_2 = 1 \}. \end{aligned}$$

The objective satisfies: $-2 z^T x \geq -2 \|z\| \|x\| = -2\|x\|$

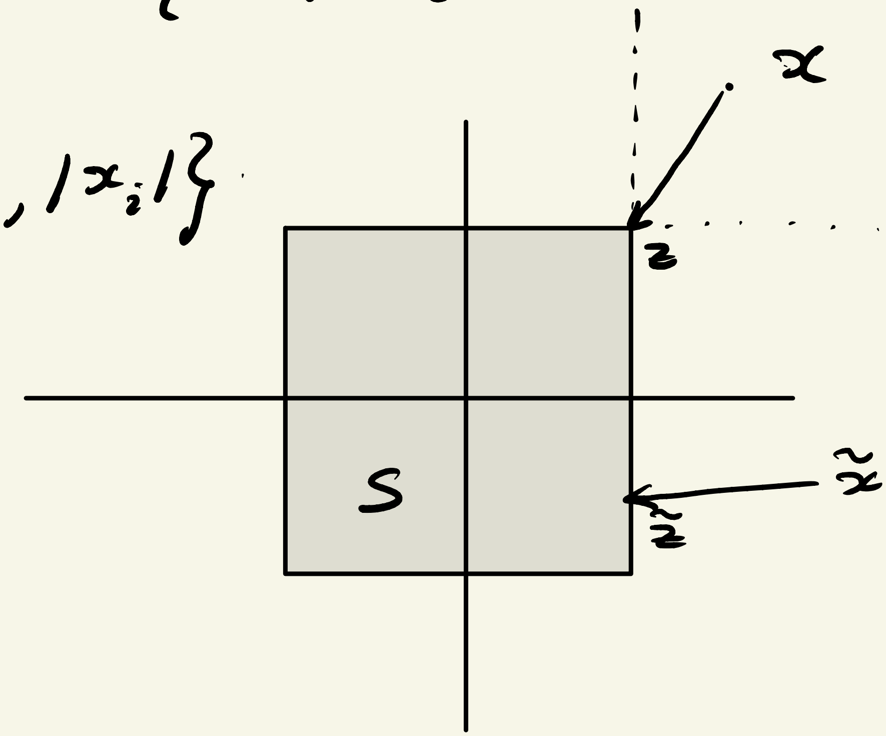
with equality if $z = x / \|x\|$.

Projection onto Inf. ball.

$$z = \text{proj}_C(x), \quad C = \{ x \mid |x_i| < 1, i = 1, \ldots n \}$$

Then,

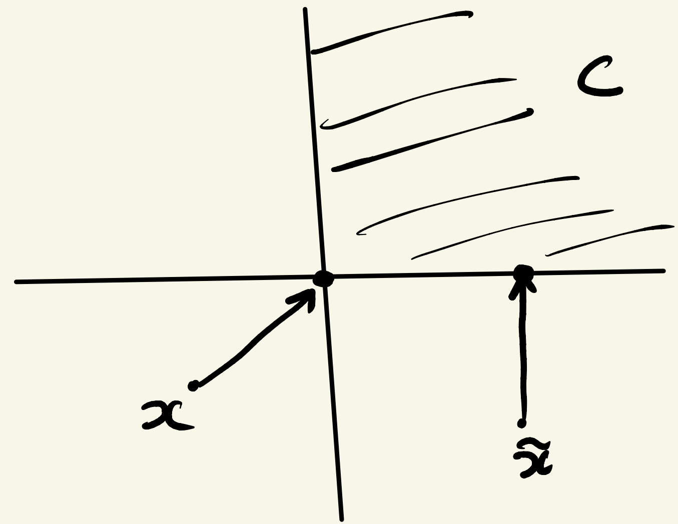
$$z_i = \text{sign}(x_i) \min\{1, |x_i|\}.$$



Projection onto nonnegative elements.

$z = \mathrm{proj}_C(x)$, $C = \{ z \mid z_i \geq 0, \ i = 1, \dots, n \}$

$z_i = \max\{0, x_i\}$



Projection onto Affine set:

$z = \text{proj}_C(x)$, $C = \{ z \mid Az = b \}$, A full row rank.

projection solves:

$$\begin{aligned} \underset{z}{\text{minimize}} \quad & \frac{1}{2} \| z - x \|_2^2 \\ \text{subject to} \quad & Az = b \end{aligned}$$

- Recall: optimality condition is

$$z^* - x \in \text{Range}(A^T) \; , \quad Az^* = b .$$

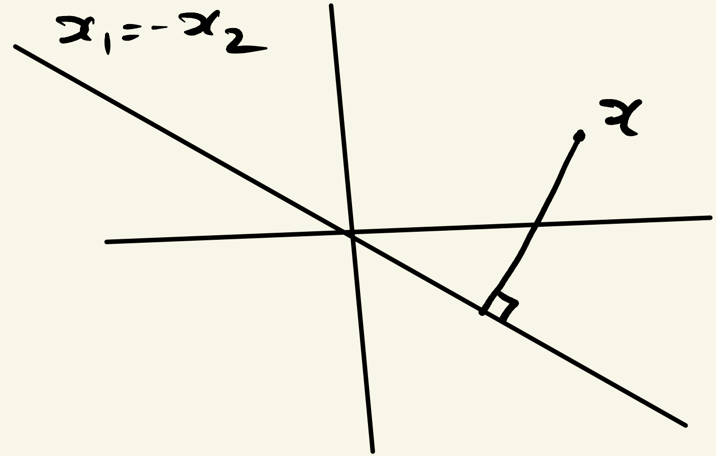
$$A^T y = z^* - x \Rightarrow AA^T y = Az^* - Ax .$$

$$\Rightarrow \quad y = (AA^T)^{-1} (b - Ax)$$

so, $z^* = A^T (AA^T)^{-1} (b - Ax) + x$

Example : De-biasing.

$$z = \text{proj}_C(x) \quad , \quad C = \{ z \mid e^T z = 0 \}$$



- Using $A = e^T$, $b = 0$

$$z^* = x + A^T \underbrace{(AA^T)}_{n}^{-1} (\underset{0}{\underset{\sim}{b}} - Ax)$$

$$= x - \frac{1}{n} \sum_{i=1}^{n} x_i$$

- z^* is de-biased x.

Projected Gradient descent

$$\min_{x} \quad f(x) \quad \text{subject to } x \in S.$$

Projected gradient descent

$$x^{(k+1)} = \mathrm{proj}_S\left(x^{(k)} - \alpha^{(k)} \nabla f(x^{(k)})\right)$$