Projection operator is non-expansive

and also firmly non-expansive

Andersen Ang

Mathématique et recherche opérationnelle, UMONS, Belgium

manshun.ang@umons.ac.be Homepage: angms.science

First draft: November 21, 2018
Last update: December 22, 2020

Euclidean projection

- Given a closed, non-empty, convex set $C \subset \mathbb{R}^n$, and a point $\mathbf{z} \in \mathbb{R}^n$ ($\mathbf{z}$ can be in or outside C), the Euclidean projection P produce a point x that
$$\mathbf{x} = P(\mathbf{z}) = \underset{\mathbf{x} \in C}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{x} - \mathbf{z}\|_2^2.$$
i.e., input of P is a point $\mathbf{z}$, the output of P is a point $\mathbf{x} \in C$ that is closest to $\mathbf{z}$.

- This problem always has a unique solution because of Weierstrass's theorem. Details here.

- Here we consider Euclidean projection. Example of other projection: Bregman projection.

Euclidean projection

$$\mathbf{x} = P(\mathbf{z}) = \operatorname*{argmin}_{\mathbf{x} \in C} \frac{1}{2}\|\mathbf{x} - \mathbf{z}\|_2^2.$$

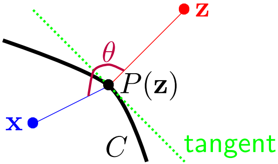
- Two possibilities between $\mathbf{z}$ and C:
  - $\mathbf{z} \in C$.
    Projection becomes *Identity operator*: it output $\mathbf{x}$ as $\mathbf{z}$
    $$P(\mathbf{z}) = \mathbf{z}.$$
  - $\mathbf{z} \notin C$.
    Projection finds the closest point $\mathbf{x} \in C$ to $\mathbf{z}$.
    - Distance here is defined in terms of Euclidean distance $\|\mathbf{v}\|_2 = \sqrt{v_1 + v_2 + \cdots + v_n}$.
    - A property of P : $\mathbf{x}$ will be located on the boundary of C: $\mathbf{x} \in$bd$C$.

Bourbaki-Cheney-Goldstein inequality[2]

- Modern names: Obtuse angle criterion, Projection theorem[1].
- What is it: a variational characterization of projection operator

$$\left\langle \mathbf{z} - P(\mathbf{z}), \mathbf{x} - P(\mathbf{z}) \right\rangle \leq 0, \quad \forall \mathbf{x} \in C.$$

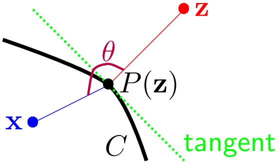
The angle in between is obtuse ($\theta \geq 90^\circ$).



[1] Note the name "projection theorem" is usually refer to projection in the context of Hilbert space (= vector space equipped with inner product, an operation that allows defining lengths and angles.)

[2] E. W. Cheney and A. A. Goldstein, Tchebycheff approximation and related extremal problems, J. Math. Mech. 14 (1965), 87-98.

Why $\left\langle \mathbf{z} - P(\mathbf{z}), \mathbf{x} - P(\mathbf{z}) \right\rangle \leq 0, \quad \forall \mathbf{x} \in C$. is true



- First, the line segment $L_\lambda : (1-\lambda)P(\mathbf{z}) + \lambda \mathbf{x}$ for $\lambda \in [0,1]$ is always inside C.

- Next, let $\mathbf{z} - L_\lambda$ be the distance between $\mathbf{z}$ and the line L_λ.

- As $L_\lambda \in C$, so we want $\mathbf{z}$ to be close to L_λ as close as possible. That is, we want to minimize distance($\mathbf{z}$, L_λ):

$$\phi(\lambda) = \text{distance}(\mathbf{z},\, L_\lambda) := \left\| \mathbf{z} - \Big((1-\lambda)P(\mathbf{z}) + \lambda \mathbf{x} \Big) \right\|_2^2 .$$

  Expand $\phi(\lambda)$ gives an quadratic function on λ. Such parabola has a vertex, minimizing ϕ at the vertex gives the obtuse angle criterion.

- We have $\phi(\lambda) = a\lambda^2 + b\lambda + c$

$$\begin{aligned}\phi(\lambda) &= \left\| \mathbf{z} - \Big((1-\lambda) P(\mathbf{z}) + \lambda \mathbf{x} \Big) \right\|_2^2 \\ &= \left\| \Big(\mathbf{z} - P(\mathbf{z}) \Big) - \lambda \Big(\mathbf{x} - P(\mathbf{z}) \Big) \right\|_2^2 \\ &= \underbrace{\|\mathbf{z} - P(\mathbf{z})\|_2^2}_{c} \underbrace{-2 \Big\langle \mathbf{z} - P(\mathbf{z}), \mathbf{x} - P(\mathbf{z}) \Big\rangle}_{b} \lambda + \underbrace{\|\mathbf{x} - P(\mathbf{z})\|_2^2}_{a} \lambda^2.\end{aligned}$$

- The vertex of $a\lambda^2 + b\lambda + c$ is at $\lambda = \dfrac{-b}{2a}$

$$\lambda = -\frac{-2\Big\langle \mathbf{z} - P(\mathbf{z}), \mathbf{x} - P(\mathbf{z}) \Big\rangle}{2\|\mathbf{x} - P(\mathbf{z})\|_2^2 2} = \frac{\Big\langle \mathbf{z} - P(\mathbf{z}), \mathbf{x} - P(\mathbf{z}) \Big\rangle}{\|\mathbf{x} - P(\mathbf{z})\|_2^2}.$$

- We want λ to be non-positive: so that on the line segment L_λ, the parabola is minimized at 0. λ is non-positive gives $\langle z - P(z), x - P(z) \rangle \leq 0$.
  (note: the inner product $\langle z - P(z), x - P(z) \leq 0$ doesn't need to be negative, only non-positive.)

Projection operator is non-expansive

- A function f is called *non-expansive*[3] if f is L-Lipschitz with $L \leq 1$. That is, for any two points $\mathbf{x}, \mathbf{z} \in \mathrm{dom}\ f$,

$$\|f(\mathbf{x}) - f(\mathbf{z})\| \leq L\|\mathbf{x} - \mathbf{z}\|, \quad \text{where } L \leq 1.$$

  It means the distance between the mapped points is smaller than that of the unmapped points in the original space.

- Projection operator is non-expansive:

$$\|P(\mathbf{x}) - P(\mathbf{z})\|_2 \leq \|\mathbf{x} - \mathbf{z}\|_2.$$

- Next slide: variational characterization implies non-expansiveness. i.e.,

$$\langle \mathbf{z} - P(\mathbf{z}), \mathbf{x} - P(\mathbf{z})\rangle \leq 0 \ \ \forall \mathbf{x} \in C \implies \|P(\mathbf{x}) - P(\mathbf{z})\|_2 \leq \|\mathbf{x} - \mathbf{z}\|_2.$$

[3] Non-expansive becomes *contractive* if $L < 1$.

Begins with the variational characterization

$$\langle \mathbf{z} - P(\mathbf{z}) \ , \ \mathbf{x} - P(\mathbf{z}) \rangle \leq 0 \ \ \forall \mathbf{x} \in C. \tag{1}$$

replace $\mathbf{x}$ by $P(\mathbf{x})$ in (1) $\qquad \langle \mathbf{z} - P(\mathbf{z}) \ , \ P(\mathbf{x}) - P(\mathbf{z}) \rangle \leq 0.$ (2)

replace $\mathbf{z}$ by $\mathbf{x}$ and $\mathbf{x}$ by $P(\mathbf{z})$ in (1) $\qquad \langle \mathbf{x} - P(\mathbf{x}) \ , \ P(\mathbf{z}) - P(\mathbf{x}) \rangle \leq 0.$ (3)

(2)+(3) will cancel $P(\mathbf{z}) - P(\mathbf{x})$, not good. So flip the sign of (3) gives

$$\langle P(\mathbf{x}) - \mathbf{x} \ , \ P(\mathbf{x}) - P(\mathbf{z}) \rangle \leq 0. \tag{4}$$

$$\begin{array}{cl}
(2) + (4) & \langle \mathbf{z} - P(\mathbf{z}) + P(\mathbf{x}) - \mathbf{x} \ , \ P(\mathbf{x}) - P(\mathbf{z}) \rangle \leq 0 \\
\iff & \langle \mathbf{z} - \mathbf{x} + P(\mathbf{x}) - P(\mathbf{z}) \ , \ P(\mathbf{x}) - P(\mathbf{z}) \rangle \leq 0 \\
\iff & \langle \mathbf{z} - \mathbf{x}, P(\mathbf{x}) - P(\mathbf{z}) \rangle \leq - \langle P(\mathbf{x}) - P(\mathbf{z}), P(\mathbf{x}) - P(\mathbf{z}) \rangle \\
\iff & \langle \mathbf{z} - \mathbf{x}, P(\mathbf{z}) - P(\mathbf{x}) \rangle \geq \| P(\mathbf{x}) - P(\mathbf{z}) \|_2^2 \\
\iff & \| (\mathbf{z} - \mathbf{x})^\top \big(P(\mathbf{z}) - P(\mathbf{x}) \big) \|_2 \geq \| P(\mathbf{x}) - P(\mathbf{z}) \|_2^2
\end{array}$$

By Cauchy-Schwarz inequality, the left-hand-side is upper bounded by $\|\mathbf{z} - \mathbf{x}\|_2 \|P(\mathbf{z}) - P(\mathbf{x})\|_2^{\frac{1}{2}}$, we get $\|\mathbf{z} - \mathbf{x}\|_2 \|P(\mathbf{z}) - P(\mathbf{x})\|_2^{\frac{1}{2}} \geq \|P(\mathbf{x}) - P(\mathbf{z})\|_2^2$. Cancels $\|P(\mathbf{x}) - P(\mathbf{z})\|_2^{\frac{1}{2}}$ finishes the proof. $\square$

Projection operator is firmly non-expansive

- In fact, projection operator is not just non-expansive, but also *firmly non-expansive* (stronger condition).

- A definition[4] of firmly non-expansive is

$$\big\langle \big(\mathbf{x} - P(\mathbf{x})\big) - \big(\mathbf{z} - P(\mathbf{z})\big)\ ,\ P(\mathbf{x}) - P(\mathbf{z})\big\rangle \geq 0.$$

  In fact, we have this already: from the proof of non-expansiveness, (2)+(4) gives

$$\langle \mathbf{z} - P(\mathbf{z}) + P(\mathbf{x}) - \mathbf{x}\ ,\ P(\mathbf{x}) - P(\mathbf{z})\rangle \leq 0,$$

  we get firmly non-expansiveness by flipping the inequality sign and the sign of the red terms.

End of document

[4] See Proposition 4.4 in the book by Bauschke and Combettes "Convex Analysis and Monotone Operator Theory in Hilbert Spaces, 2nd edition"