Projection onto nonnegative orthant, rectangular box and polyhedron

Andersen Ang

ECS, Uni. Southampton, UK

andersen.ang@soton.ac.uk

Homepage angms.science

Version: April 2, 2023
First draft: March 19, 2020

Content

Projection onto nonnegative orthant

- The Euclidean projection of a given point $\boldsymbol{y} \in \mathbb{R}^n$ onto a (non-empty & compact) set $S \subseteq \mathbb{R}^n$, denoted as $\operatorname{proj}_S(\boldsymbol{y})$, is a function $\mathbb{R}^n \to \mathbb{R}^n$ that output a point $\hat{\boldsymbol{x}}$ by solving

  $$\hat{\boldsymbol{x}} = \operatorname{proj}_S(\boldsymbol{y}) = \operatorname*{argmin}_{\boldsymbol{x} \in S} \|\boldsymbol{x} - \boldsymbol{y}\|_2.$$

  Such optimization always has a unique solution, details here.

- **Question**: What if S is the nonnegative orthant?

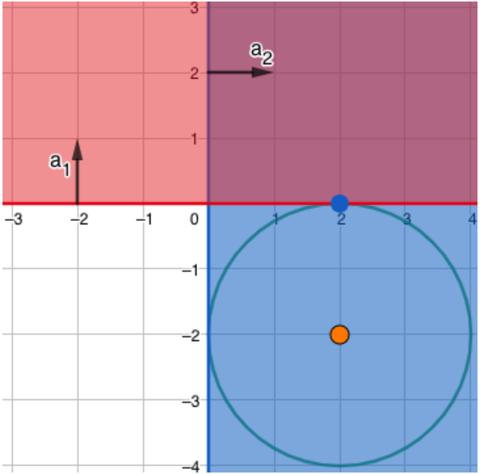
  $$\operatorname{proj}_S(\boldsymbol{y}) = \operatorname*{argmin}_{\boldsymbol{x} \geq 0} \|\boldsymbol{x} - \boldsymbol{y}\|_2,$$

  where $\boldsymbol{x} \geq 0$ means $\boldsymbol{x}$ is inside the nonnegative orthant $S = \{\, \boldsymbol{x} \,|\, x_i \geq 0 \ \forall i \,\}$.

- This problem has a trivial solution $\hat{\boldsymbol{x}} = [\boldsymbol{y}]_+$, where $\hat{x}_i = [y_i]_+$ with $[\,\cdot\,]_+ = \max\{\,\cdot\,, 0\,\}$.

$$[\boldsymbol{y}]_+ = \operatorname*{argmin}_{\boldsymbol{x}\geq 0} \|\boldsymbol{x}-\boldsymbol{y}\|_2$$

Among all the feasible value in the nonnegative orthant, 0 is the closest to negative value.



Example

- The orange point $\boldsymbol{y} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$ is projected to the blue point $\begin{bmatrix} 2 \\ 0 \end{bmatrix}$ as the blue point is where the blue circle just touch the nonnegative orthant (violet region).
- Here the radius of the circle (=2) is the optimal cost value.

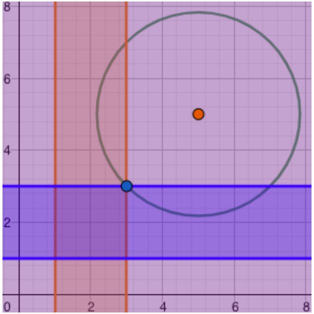
Projection onto a box

A generalization of the nonnegativity is the box constraint

$$P_S(\boldsymbol{y}) = \underset{\boldsymbol{l} \le \boldsymbol{x} \le \boldsymbol{u}}{\operatorname{argmin}} \|\boldsymbol{x} - \boldsymbol{y}\|_2,$$

with the close form solution $[P_S(\boldsymbol{y})]_i = \begin{cases} l_i & y_i \le l_i \\ y_i & l_i < y_i < u_i \\ u_i & y_i \ge u_i \end{cases}$

An example $l_i = 1, u_i = 2$ and a point at $(5, 5)$.



Projection onto a polyhedron

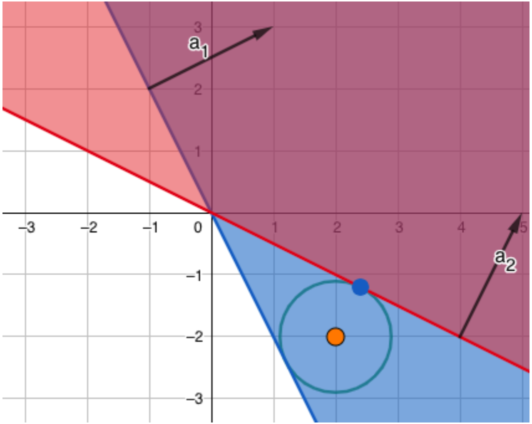
- $\operatorname{proj}_{\text{nonnegative orthant}}$ and $\operatorname{proj}_{\text{box}}$ are special cases of $\operatorname{proj}_{\text{polyhedron}}$

$$\text{Polyhedron} \;\; S = \Big\{ \boldsymbol{x} \mid \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b} \Big\}.$$

  where $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ is full rank, and $\boldsymbol{b} \in \mathbb{R}^m$.

- Note that it does not matter $\boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b}$ or $\boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}$: we can absorb the negative sign into $\boldsymbol{A}$ or $\boldsymbol{b}$.

Example 1 : $m = n = 2$



orange point	$\boldsymbol{y} = [2, -2]$
blue point	$\hat{\boldsymbol{x}} = [12/5, -6/5]$
Red region	$x + 2y \geq 0$
Blue region	$2x + y \geq 0$
$\boldsymbol{a}_1$	$[2, 1]$
$\boldsymbol{a}_2$	$[1, 2]$

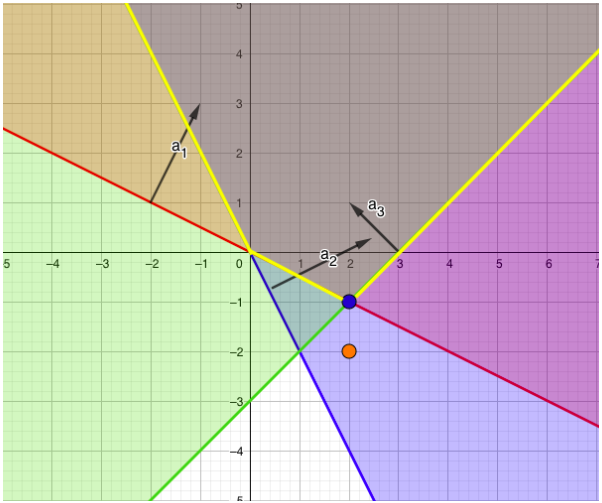
$$\boldsymbol{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \quad \boldsymbol{b} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

- $\boldsymbol{y}$ fulfill $\boldsymbol{Ax} \geq \boldsymbol{b}$ for the first row of $\boldsymbol{A}$, but not the second row.
- The optimal cost value $= \frac{4}{5}$, which is the distance between $\hat{\boldsymbol{x}}$ and $\boldsymbol{y}$, and it is also the radius of the blue circle.
- We can see in this example that, proj to nonnegative orthant is simply the vector $\boldsymbol{a}_1, \boldsymbol{a}_2$ are rotated to the direction of the standard basis.

Example 2 : $m = 3, n = 2$

orange point	$\boldsymbol{y} = [2, -2]$
blue point	$\hat{\boldsymbol{x}} = [2, -1]$
Red region	$x + 2y \geq 0$
Blue region	$2x + y \geq 0$
Green region	$x - y \leq 3$

$$\boldsymbol{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \\ -1 & 1 \end{bmatrix} \quad \boldsymbol{b} = \begin{bmatrix} 0 \\ 0 \\ -3 \end{bmatrix}$$



- The boundary of the feasible region is highlighted in yellow.

Solving projection onto polyhedron

- Unlike the projection onto nonnegative orthant or projection onto a box, the problem

$$\underset{\boldsymbol{A}\boldsymbol{x}\leq\boldsymbol{b}}{\operatorname{argmin}}\,\|\boldsymbol{x}-\boldsymbol{y}\|_2,$$

  has no simple close form solution in general.

- To solve it, we use an iterative optimization algorithms, such as
  1. Alternating projection algorithm
  2. Active set methods
  3. Interior point methods
  4. (Dual) Proximal gradient method
  5. Douglas-Rachford splitting algorithm on the dual

  We talk very briefly on the first four approaches.

Alternating projection algorithm

- The constraint $\boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}$ describe a (non-empty) polyhedron as an intersection of m inequality constraint, each of the constraint is a problem pf projection onto halfspace

$$\operatorname*{argmin}_{\langle \boldsymbol{a}, \boldsymbol{x} \rangle \leq b} \frac{1}{2} \|\boldsymbol{x} - \boldsymbol{y}\|_2^2,$$

with the close form solution as

$$\hat{\boldsymbol{x}} = \begin{cases} \boldsymbol{y} & \langle \boldsymbol{a}, \boldsymbol{x} \rangle \leq b \\ \boldsymbol{y} - \dfrac{\langle \boldsymbol{a}, \boldsymbol{x} \rangle - b}{\|\boldsymbol{a}\|_2^2} \boldsymbol{a} & \langle \boldsymbol{a}, \boldsymbol{x} \rangle > b \end{cases}$$

- Alternating projection algorithm: cycle through the projection onto halfspace from $i = 1$ to r and repeats a few times gives the projection onto $\boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}$.

- Drawback of this approach: possibly slow if m is large.

Active set

- Assume we know a set $\mathcal{I} : \{i \in [m] \mid \langle \boldsymbol{a}^i, \boldsymbol{x} \rangle = b_i\}$[1] that at the solution the equality is satisfied. Let $\boldsymbol{A}_{\mathcal{I}}$ be the submatrix of $\boldsymbol{A}$ with rows in $\mathcal{I}$. Then we can solve the following problem instead
$$\operatorname*{argmin}_{\boldsymbol{x}} \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{y}\|_2^2 \text{ s.t. } \boldsymbol{A}_{\mathcal{I}}\boldsymbol{x} = \boldsymbol{b}_{\mathcal{I}}.$$
The optimality conditions can be expressed as a matrix-vector equation as
$$\begin{bmatrix} \boldsymbol{I} & \boldsymbol{A}_{\mathcal{I}}^\top \\ \boldsymbol{A}_{\mathcal{I}} & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{z} \end{bmatrix} = \begin{bmatrix} 0 \\ \boldsymbol{b}_{\mathcal{I}} \end{bmatrix}$$
- The idea of active set method is to construct a set $\mathcal{J}$ to approximate $\mathcal{I}$. In each iteration, it solves the above equation, and correct $\mathcal{J}$. Eventually $\mathcal{J}$ approach to $\mathcal{I}$ and the sol. of the matrix-vector equation gives the sol. of the original problem.

[1] $[n]$ means the set $\{1, 2, \cdots, n\}$ and $\boldsymbol{a}^i$ is the i-th row of $\boldsymbol{A}$

Interior point method

- First, transform $\boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}$ to $\boldsymbol{A}\boldsymbol{x} + \boldsymbol{s} = \boldsymbol{b}$, where $\boldsymbol{s} \geq 0$ is a slack variable.

- Construct a log barrier on $\boldsymbol{s}$ with parameter $\mu > 0$, and put it into the cost function, we obtain a typical problem in the interior point method

$$\underset{\boldsymbol{x}}{\operatorname{argmin}} \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{y}\|_2^2 - \mu \sum_i \log|s_i| \text{ s.t. } \boldsymbol{A}\boldsymbol{x} + \boldsymbol{s} = \boldsymbol{b}.$$

  This problem is convex.

- Starting from $(\boldsymbol{x}_0, \mu_0)$, compute an approximate solution $\boldsymbol{x}$, then decrease μ. As $\mu \to 0$, $\boldsymbol{x}_k \to \boldsymbol{x}^*$.

- Drawback of interior point method: not scalable, not for problem with large size.

(Dual) proximal gradient method

- The dual of the problem is

$$\min_{\boldsymbol{z}} \frac{1}{2}\|\boldsymbol{A}^\top \boldsymbol{z} - \boldsymbol{A}^\top \boldsymbol{y}\|_2^2 + \langle \boldsymbol{b}, \boldsymbol{z}\rangle \ \text{ s.t. } \ \boldsymbol{z} \geq 0.$$

  The key is that nonnegative constraint is easy to handle, so you keep the nonnegative constraint.
- Proximal update of this problem is simple. See here for details of (unaccelerated) proximal gradient.
- In fact, dual proximal gradient method in this case is exactly the dual projected gradient method.
- Nesterov's acceleration can be used.
- See here for the discussion of dual proximal gradient method, and also how the dual problem is constructed.

Last page - summary

Discussed: solving projection onto a polyhedron

- Polyhedron set : $\boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}$, $\boldsymbol{A}$ full rank
- Brief overview of solving projection problem by alternating projection algorithm, active set method, interior point method and dual proximal gradient method.
- Projection onto nonnegative orthant and box as special cases with simple close form solution.

Not discussed

- If $\boldsymbol{A}$ not full rank.
- Details of the methods.
- Solving the projection problem using Douglas-Rachford splitting.

End of document