Projection onto nonnegative orthant, rectangular box and polyhedron, and

- alternating projection algorithm
- active set method
- interior point method
- dual proximal gradient method.

Andersen Ang

Mathématique et recherche opérationnelle, UMONS, Belgium

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

First draft: March 19, 2020
Last update: December 23, 2020

Projection onto nonnegative orthant

- Given $\mathbf{y} \in \mathbb{R}^n$, the Euclidean projection of $\mathbf{y}$ onto a (non-empty and compact)set $S \subseteq \mathbb{R}^n$, denoted as $P_S(\mathbf{y})$, is a function $\mathbb{R}^n \to \mathbb{R}^n$ that output a point $\hat{\mathbf{x}}$ by solving the following optimization problem

$$\hat{\mathbf{x}} = P_S(\mathbf{y}) = \underset{\mathbf{x} \in S}{\operatorname{argmin}} \|\mathbf{x} - \mathbf{y}\|_2.$$

  Such optimization always has a unique solution, details here.

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

$$P_S(\mathbf{y}) = \underset{\mathbf{x} \geq 0}{\operatorname{argmin}} \|\mathbf{x} - \mathbf{y}\|_2,$$

  where $\mathbf{x} \geq 0$ means $\mathbf{x}$ is inside the set

$$\text{Nonnegative orthant} \quad S = \{\mathbf{x} \mid x_i \geq 0, \forall i\}.$$

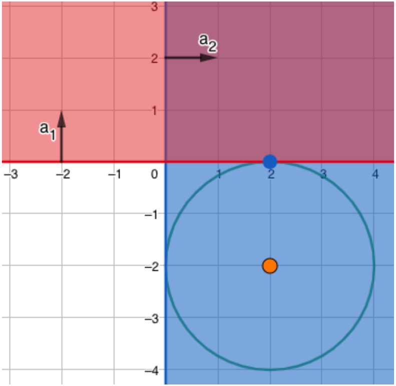
- This problem has a trivial solution $\hat{\mathbf{x}} = [\mathbf{y}]_+$, where $\hat{x}_i = [y_i]_+$ with $[\ \cdot\ ]_+ = \max\{\,\cdot\,, 0\,\}$.

$[\mathbf{y}]_+$ is the solution of $\underset{\mathbf{x}\geq 0}{\operatorname{argmin}} \|\mathbf{x}-\mathbf{y}\|_2$

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

- An example

  The orange point $\mathbf{y} = [2, -2]^\top$ is projected to the blue point $[2, 0]$ as the blue point is where the blue circle just touch the nonnegative orthant (the region covered by both red and blue color). Here the radius of the circle $(=2)$ is the optimal cost value.



Projection onto a box

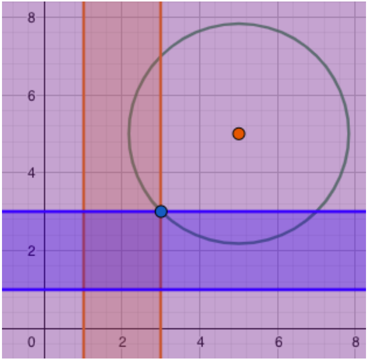
- A generalization of the nonnegativity is the box constraint

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

with the close form solution

$$[P_S(\mathbf{y})]_i = \begin{cases} l_i & y_i \leq l_i \\ y_i & l_i < y_i < u_i \\ u_i & y_i \geq u_i \end{cases}$$

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



Projection onto a polyhedron

- Projection onto nonnegative orthant and projection onto a box are special cases to projection onto a polyhedron

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

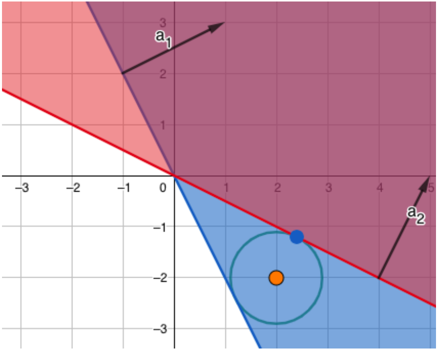
  where $\mathbf{A} \in \mathbb{R}^{m \times n}$ is full rank, and $\mathbf{b} \in \mathbb{R}^m$.

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

Example 1 : $m = n = 2$

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

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

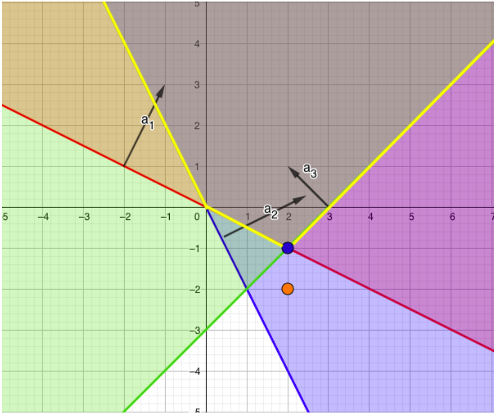


- $\mathbf{y}$ fulfill $\mathbf{Ax} \geq \mathbf{b}$ for the first row of $\mathbf{A}$, but not the second row.
- The optimal cost value $= \frac{4}{5}$, which is the distance between $\hat{\mathbf{x}}$ and $\mathbf{y}$, and it is also the radius of the blue circle.
- We can see in this example that,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

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

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

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



- 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

$$\operatorname*{argmin}_{\mathbf{A}\mathbf{x}\leq\mathbf{b}} \|\mathbf{x}-\mathbf{y}\|_2,$$

  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. We do not talk about approach 5 here.

Alternating projection algorithm

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

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

with the close form solution as

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

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

- Drawback of this approach: if m is large, the cycle takes time.

Active set

- As the cost function is strongly convex and the constraint set is non-empty and compact (and also convex), so there exists a unique global minimizer.
- Assume we know beforehand that, the set i of the solution such that equality is satisfied in the i^{th} constraint. Mathematically[1], $\mathcal{I} : \{i \in [m] \mid \langle \mathbf{a}^i, \mathbf{x} \rangle = b_i\}$. Let $\mathbf{A}_{\mathcal{I}}$ be the submatrix of $\mathbf{A}$ with rows in $\mathcal{I}$. Then we can solve the following problem instead
$$\operatorname*{argmin}_{\mathbf{x}} \frac{1}{2}\|\mathbf{x} - \mathbf{y}\|_2^2 \quad \text{s.t.} \quad \mathbf{A}_{\mathcal{I}}\mathbf{x} = \mathbf{b}_{\mathcal{I}}.$$
The optimality conditions can be expressed as a matrix-vector equation as
$$\begin{bmatrix} \mathbf{I} & \mathbf{A}_{\mathcal{I}}^\top \\ \mathbf{A}_{\mathcal{I}} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{z} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{b}_{\mathcal{I}} \end{bmatrix}$$
- 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 $\mathbf{a}^i$ is the i-th row of $\mathbf{A}$

Interior point method

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

- Construct a log barrier on $\mathbf{s}$ with parameter $\mu > 0$, and put it into the cost function, we obtain the typical problem setup in the interior point method
$$\underset{\mathbf{x}}{\operatorname{argmin}} \ \frac{1}{2}\|\mathbf{x} - \mathbf{y}\|_2^2 - \mu \sum_i \log |s_i| \ \text{s.t.} \ \mathbf{A}\mathbf{x} + \mathbf{s} = \mathbf{b}.$$
This problem is convex.

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

- Drawback of interior point method: non-scalable, it cannot handle problem with very large size.

(Dual) proximal gradient method

- The dual of the problem is

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

  The key is that nonnegative constraint is easy to handle by proximal operator, so you keep the nonnegative constraint
- Proximal update of this problem is simple: the cost function is a quadratic function. The standard gradient descent approach works here : take the gradient, take the step size as $\frac{1}{L}$ for L as the Lipschitz constant of the gradient, perform update, project. 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 : $\mathbf{Ax} \leq \mathbf{b}$, $\mathbf{A}$ 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 $\mathbf{A}$ is not full rank.
- Details of the methods.
- Solving the projection problem using Douglas-Rachford splitting.

End of document