Solving piece-wise linear equation

Andersen Ang

Mathématique et recherche opérationnelle
UMONS, Belgium

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

First draft : October 8, 2020
Last update : October 8, 2020

Piece-wise linear function

- A piece-wise linear function

$$f(x) = \sum_{i=1}^{n} c_i \max\left\{0, a_i x + b_i\right\} - d,$$

where a_i, b_i, c_i and d are given real numbers.

- Using $[\,\cdot\,]_+$ to denote $\max\{\,0\,,\,\cdot\,\}$. A piece-wise linear equation is defined as

$$\sum_{i=1}^{n} c_i \big[a_i x + b_i\big]_+ = d.$$

- This document: discuss how to solve this equation, i.e., solve for the root x.

Where do you find piece-wise linear function

- Discrete optimization.
  Piece-wise linear function appears naturally in many discrete optimization problems.

- Projection onto simplex.
  Projecting a vector in $\mathbb{R}^n$ onto the convex set

  $$\{\mathbf{x} \mid \mathbf{x} \geq 0, \langle \mathbf{x}, \mathbf{b} \rangle = 1\}$$

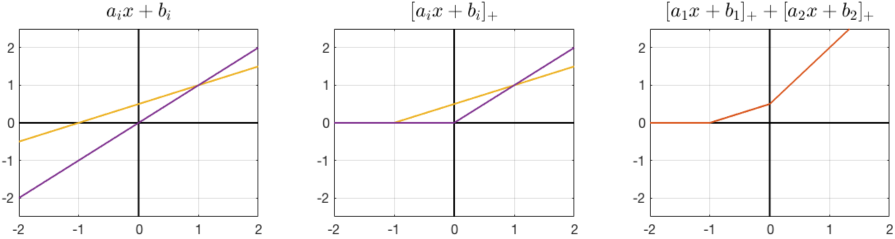
  involves solving a piece-wise linear equation.

Intuition ... (1/2)

- For simplicity, consider the special case with $c_i = 1$ and $n = 2$:

$$\left[0, a_1 x + b_1\right] + \left[0, a_2 x + b_2\right] = d.$$

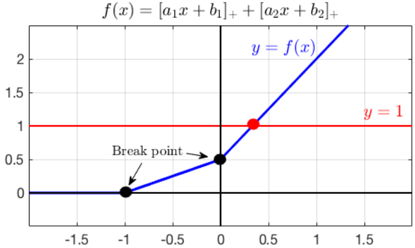
Say $a_1 > 0, a_2 > 0,\ b_1 > 0$ and $b_2 = 0$



For $d = 0$, we see the solution set is $\{x \mid x \leq -1\}$.

Intuition ... (2/2)

For $d = 1$, solving $\big[0, a_1 x + b_1\big] + \big[0, a_2 x + b_2\big] = 1$ means



to find the red point.

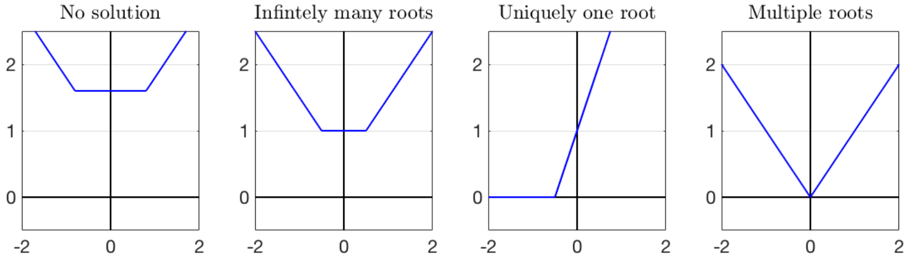
From both $d = 0$ and $d = 1$, we see that

- The piece-wise linear function f is divided into 3 line segments by 2 break points.
- Finding the root corresponds to finding the point that intersects f and the line $y = d$. The x-coordinate of the point will be the solution.
- This gives us a solution approach: find the line segment that intersects with $y = d$.

Nature of the solution

- In general $f(x)$ can have
  - unique root.
  - multiple roots.
  - infinitely many roots.
  - no solution.

- For examples: $\left[a_1 x + b_1\right]_+ + \left[a_2 x + b_2\right]_+ = 1,$



Solution approach

- We now consider the solution approach of a particular problem.
- Fist we simplify the problem.
- There are in general two cases: if $d = 0$, we have

  $$\sum_{i=1}^{n} c_i \big[a_i x + b_i\big]_+ = 0.$$

  If $d \neq 0$, we divide the equation by d. Absorbing d into c_i gives

  $$\sum_{i=1}^{n} c_i \big[a_i x + b_i\big]_+ = 1.$$

- We focus on the second problem as it is harder.

Solution approach ... (1/2)

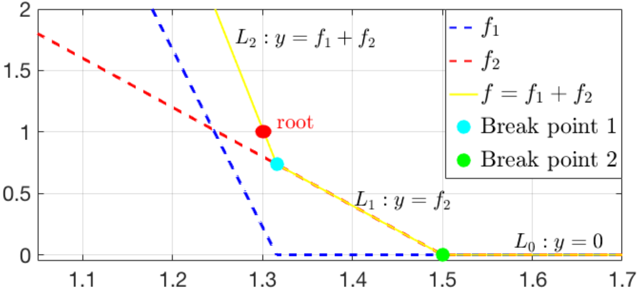
For understanding, we focus on solving an example:

$$\underbrace{3.8\big[-3.8x+5\big]_+}_{f_1}+\underbrace{2\big[-2x+3\big]_+}_{f_2}=1.$$

The break point (the point that changes the behavior of the function) of f_1, f_2 are:

$$BP_1=\frac{5}{3.8}=1.3158, \quad BP_2=\frac{3}{2}=1.5.$$

The break points divide the f into three line segments: L_0, L_1 and L_2.



Solution approach ... (2/2)

- Firs we test L_0: does $y = 1$?
  No since $y = 0$ for all $x \in \mathrm{dom} L_0 = \{w \mid w \geq 1.5\}$.

- Then we move to L_1: does $y = 1$?
  Np since $y = f_2 < 1$ for $x \in \mathrm{dom} L_1 = \{w \mid 1.3158 \leq w \leq 1.5\}$

- Then we move to L_2: does $y = 1$?
  Yes since $y = f_1 + f_2$ contains 1 for $x \in \mathrm{dom} L_2 = \{w \mid w \leq 1.3158\}$.
  And the root x can be computed as

$$
\begin{array}{rrcl}
 & f(x) & = & 1 \\
\iff & f_1(x) + f_2(x) & = & 1 \\
\iff & 3.8[-3.8x+5]_+ + 2[-2x+3]_+ & = & 1 \\
\iff & 3.8(-3.8x+5) + 2(-2x+3) & = & 1 \\
\iff & -18.4x + 25 & = & 1 \\
\iff & x & = & 1.304.
\end{array}
$$

  The 3rd $\iff$ is the most important one: since f_1 and f_2 are nonnegative in $\mathrm{dom} L_2$, we can remove $[\,\cdot\,]_+$.

Systemic solution approach

From the previous example, we arrive at a systemic approach to solve piece-wise linear equation

- Sort the break points: this gives us the interval for each of the line segments L_i.
- Starting from the right most line segment L_0 (correspond to the largest break point):
  - Compute the equation of the line segment L_i in the form of $y = mx + c$.
  - Check whether L_i intersect the line $y = 1$.
  - If L_i does not contains the solution, move to L_{i+1}.
- Note that for the problem with $d = 0$, the line segment L_0 is the solution.
- For the problem with $d \neq 0$, the line segment L_0 definitely does not contains the solution and hence can be ignored.

The complexity

- For the general problem

$$f(x) = \sum_{i=1}^{n} c_i \big[a_i x + b_i\big]_+ = d,$$

we are given n sets of points (a_i, b_i, c_i), hence the sorting on $\dfrac{-a_i}{b_i}$ takes in general $\mathcal{O}(n \log n)$ computational cost.

- As the sorting is the most expensive step in the algorithm, the whole algorithm to solve the piece-wise linear equation thus has the general complexity of $\mathcal{O}(n \log n)$.

The complexity: practical remark

- Note that only the terms with $c_i \neq 0$ contributes to the sum in

$$f(x) = \sum_{i=1}^{n} c_i \left[a_i x + b_i\right]_+ = d.$$

  Hence, we can actually ignore the terms with $c_i = 0$.

- Assuming there are $K \leq n$ nonzero c_i, the sorting thus has $\mathcal{O}(n + K \log n)$ cost.

- Therefore, in general, the cost of solving the piece-wise linear equation is in between $\mathcal{O}(n + K \log n)$ and $\mathcal{O}(n \log n)$, depending on K

- For K that is much smaller than n, then $n + K \log n \approx n$ and hence we can say that the cost of solving the piece-wise linear equation is in between $\mathcal{O}(n)$ and $\mathcal{O}(n \log n)$.

End of document