

# A Note on Weighted Fermat Problem in $\mathcal{L}^p$ Space

Shikun Liu

## Abstract

In this paper, we begin by introducing a well-known geometry concept: the Fermat point in a triangle. Then, we generalize the problem and propose an iterative algorithm based on gradient descent to the weighted form in  $\mathcal{L}^p$  space. We also build specific solutions to some special norms: one, two and infinity. We show that the solution may not be unique in norm-one and infinity. Finally, we provide our qualitative results in the PASCAL-based program.

**Keywords:**  $\mathcal{L}^p$  space, convex optimization, weighted Fermat problem

## 1 Introduction

The Fermat point optimization problem was named after the French mathematician *Pierre de Fermat* in a private letter in 17th century: Given three points in the plane, finding the fourth point such that the sum of the distances to the three given points is the minimum. The problem was generalized by *Simpson*<sup>[1]</sup> to ask finding the minimum weighted sum of distances from three given points.

In this note, we consider the problem in  $\mathcal{L}^p$  space. Let  $\mathbf{y}_i = (y_1^i, y_2^i, \dots, y_n^i)$  be fixed points with their positive weights  $k_i$  for  $i = 1, 2, \dots, m$  for total  $m$  points. Given the optimized point  $\mathbf{F} = (x_1, x_2, \dots, x_n)$  in  $\mathbb{R}^n$ , we want to find  $(x_1, x_2, \dots, x_n)$  such that

$$\sum_{i=1}^m \left( \sum_{j=1}^n k_i |y_j^i - x_j|^p \right)^{1/p} = \min.$$

## 2 Fermat Point in Manhattan Norm

In Manhattan norm (norm-1 space), the distance equation is defined as:

$$\mathcal{L}^1 = \|\mathbf{x} - \mathbf{y}\|_1 = \sum_{i=1}^n |x_i - y_i|.$$

According to the definition of Manhattan norm, we propose our target function,

$$f(x_1, x_2, \dots, x_n) = \sum_{i=1}^m \sum_{j=1}^n k_i |y_j^i - x_j|.$$

Reorder coordinate sets  $\{y_j^i\}$  in each  $j^{\text{th}}$  dimension an increasing order  $\{\tilde{y}_j^i\}$  with its weight  $\tilde{k}_j^i$ . Define,

$$\tilde{f}_j(x_j) = \sum_{i=1}^m \tilde{k}_j^i |\tilde{y}_j^i - x_j|, \quad j = 1, 2, \dots, n$$

Since the function  $\tilde{f}_j(x_j)$  are independent in each dimension, the optimal point can be found when the derivative in each dimension of  $\tilde{f}_j(x_j)$  is zero or NOT exist.

Let  $\tilde{m}_{l_j}^i < x_j < \tilde{m}_{l_j+1}^i, l_j = 1, 2, \dots, m-1$ . Then,  $\tilde{f}_j(x_j)$  can be split into,

$$\tilde{f}_j(x_j) = \sum_{i=1}^{l_j} \tilde{k}_j^i |\tilde{y}_j^i - x_j| + \sum_{i=l_j+1}^m \tilde{k}_j^i |x_j - \tilde{y}_j^i|.$$

Therefore, the optimal point  $\mathbf{F}(x_1, x_2, \dots, x_n)$  holds when the function satisfies,

$$\frac{\partial \tilde{f}_j^-}{\partial x_j^-} = \sum_{i=1}^{l_j} \tilde{k}_j^i - \sum_{i=l_j+1}^m \tilde{k}_j^i \leq 0, \quad \frac{\partial \tilde{f}_j^+}{\partial x_j^+} = \sum_{i=1}^{l_j} \tilde{k}_j^i - \sum_{i=l_j+1}^m \tilde{k}_j^i \geq 0, \quad j = 1, 2, \dots, n$$

From the solutions above, we can see that Fermat point in norm-1 space may have infinity solutions. The visual representations with random 50 points in  $\mathbb{R}^2$  space can be illustrated as follow.

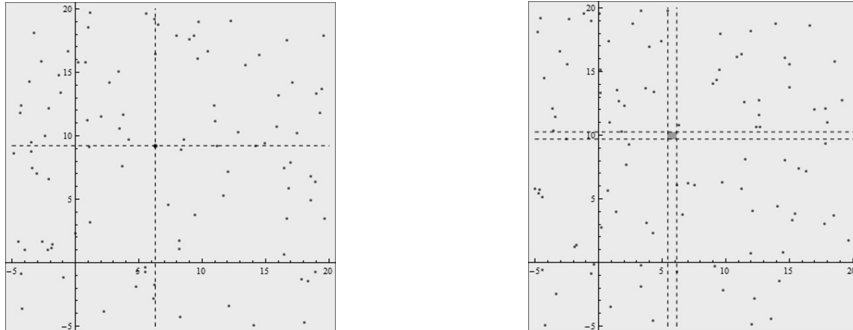


Figure 1: Two possible cases of Fermat point(s) (marked in grey point or box) in  $\mathcal{L}^1$  norm, in which left has one unique solution and right has infinity number of solutions.

### 3 Fermat Point in Euclidean Norm

The unweighted 3-points problem in the Euclidean Norm is the original Fermat Problem. A solution to the one is to construct equilateral triangles on the sides with the vertices pointing outward. The intersected vertex from three circles is the Fermat point. The visualization is Figure 2, in which the black points are fixed points, and the grey point is the Fermat Point.

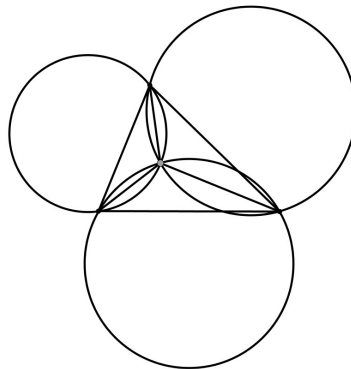


Figure 2: Fermat point in a triangle

#### 3.1 The Varignon Frame

A good tool that allow us to inspire the general Fermat problem in Euclidean space is a mechanical analogue device *The Varignon Frame* named after the French mathematician *Varignon*. A board is drilled with  $m$  holes corresponding to the coordinates of the  $m$  fixed points. One string is fed through each hole and all strings are tied together in one knot. Below the board, weights are attached to the strings and proportional to the board. Given only gravity we concern and we ignore friction, the knot will settle at the optimal point. A graph of the Varignon frame is shown in Figure 3.<sup>[2]</sup>

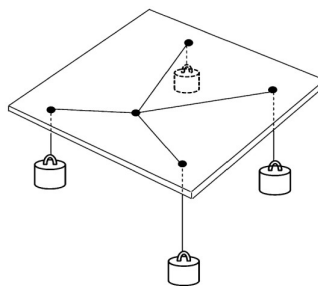


Figure 3: The Varignon Frame

**Lemma 3.1.** *The minimal potential energy point is the optimal point.*

*Proof:* Let the zero potential energy level is the ground, and height of this system is  $h$ , the length of each string is  $l_i$ , and its weight is  $G_i$ , where  $i = 1, 2, \dots, m$ . We set up an initial point  $P$  in this system, then we have

$$E(P) = \sum_{i=1}^m G_i [h - (l_i - \overline{PM}_i)] = \sum_{i=1}^m G_i (h - l_i) + \sum_{i=1}^m G_i (\overline{PM}_i)$$

When  $P$  is the minimal potential energy point, that is  $E(P)$  is the minimum,  $\sum_{i=1}^m G_i (\overline{PM}_i)$  reaches the minimum as  $\sum_{i=1}^m G_i (h - l_i)$  is the constant.  $\square$

### 3.2 The Weiszfeld's Algorithm

The Weiszfeld's iterative algorithm<sup>[3]</sup> was proposed to the Fermat problem in  $\mathbb{R}^2$  space. The idea behind this algorithm, starting anywhere and trying to converge to  $\mathbf{F}$ , is actually very simple. Let

$$\mathbf{F} = \sum_{i=1}^m \left( \sum_{j=1}^n k_j |y_j^i - F|^2 \right)^{1/2}.$$

We assume  $\mathbf{F} \notin \mathbf{y}_i$ , then the negative of the gradient of  $f$  at  $\mathbf{F}$  equals,

$$R(\mathbf{F}) = \sum_{i=1}^m k_i \frac{\mathbf{y}_i - \mathbf{F}}{\|\mathbf{y}_i - \mathbf{F}\|}.$$

Therefore, let  $R(\mathbf{F}) = 0$  which is equivalent to,

$$T(\mathbf{F}) = \sum_{i=1}^m \frac{k_i \mathbf{y}_i}{\|\mathbf{F} - \mathbf{y}_i\|} / \sum_{i=1}^m \frac{k_i}{\|\mathbf{F} - \mathbf{y}_i\|},$$

as our iterative algorithm.

*Rautenbach*<sup>[7]</sup> proposed an extension  $T^*$  of  $T$  to  $\mathbb{R}^n$  space. Let  $P \in \mathbb{R}^n$ ,

$$\left\{ \begin{array}{l} \text{If } T^*(P) = P, \text{ if and only if } P = \mathbf{F} \quad (i) \\ \text{If } T^*(P) \neq P, \text{ then } f(T^*(P)) < f(P) \quad (ii) \end{array} \right.$$

*Kuhn*<sup>[4]</sup> proved the part (ii) and showed that, given an initial point  $X_0 \in \mathbb{R}^n$ , the sequence of points  $\{X_n\}_{n \in \mathbb{N}}$  generated by the iterative process  $X_{k+1} = T(X_k)$  converges to the optimal solution of the Fermat problem if no point in the sequence is the vertex. In

this last case, Kuhn claimed to have proved that this fact only occurs for a denumerable set of initial points. This result was modified by *Chandrasekaran and Tamir* <sup>[5]</sup>, who gave two counterexamples and established the following conjecture: "If the convex hull of the set of vertices is of full dimension, then the set of initial points for which the sequence generated by the Weiszfeld's algorithm yields in a vertex is denumerable."

*Vardi and Zhang* <sup>[6]</sup> improved the limit of Weiszfeld's algorithm in case of  $\mathbf{F} \in \mathbf{y}_i$  appears in any iteration.

## 4 Fermat Point in Chebyshev Norm

The two dimensional Chebyshev distance has the similar property as the Manhattan distance. Hence, it can easily be viewed as the transformation of Manhattan distance, with the transformation matrix is

$$T = \begin{bmatrix} \frac{\sqrt{2}}{2} & 0 \\ 0 & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} \cos \frac{\pi}{2} & -\sin \frac{\pi}{2} \\ \sin \frac{\pi}{2} & \cos \frac{\pi}{2} \end{bmatrix}.$$

**Lemma 4.1.**  $\mathcal{L}^\infty(\mathbf{y}^i, \mathbf{F}) = \mathcal{L}^1(T(\mathbf{y}^i), T(\mathbf{F}))$

*Proof.*

$$\begin{aligned} \mathcal{L}^1(T(\mathbf{y}_1^i), T(\mathbf{F})) &= \frac{\sqrt{2}}{2} \left| y_1^i \cos \frac{\pi}{2} + y_2^i \sin \frac{\pi}{2} - x_1 \sin \frac{\pi}{2} - x_2 \sin \frac{\pi}{2} \right| \\ &\quad + \frac{\sqrt{2}}{2} \left| -y_1^i \cos \frac{\pi}{2} + y_2^i \sin \frac{\pi}{2} + x_1 \sin \frac{\pi}{2} - x_2 \sin \frac{\pi}{2} \right| \\ &= \frac{\sqrt{2}}{2} \max \left\{ \frac{1}{\sqrt{2}} \max \{ (y_2^i - x_2 + y_2^i - x_2), (y_1^i - x_1 + y_1^i - x_1), \right. \\ &\quad \left. (-y_1^i + x_1 - y_1^i + x_1), (-y_2^i + x_2 - y_2^i + x_2) \} \right\} \\ &= \max \{ |y_1^i - x_1|, |y_2^i - x_2| \}. \end{aligned}$$

□

However, the equivalence between  $\mathcal{L}^1$  and  $\mathcal{L}^\infty$  cannot be generalized to the  $\mathbb{R}^n$  due to the different properties on these two norms in higher dimensions.

## 5 Fermat Point in $P$ -Norm

The  $\mathcal{L}^p$  Space are the certain vector spaces of measurable functions. A vector space on which a norm is defined is called a normed vector space. It satisfies the following conditions,<sup>[8]</sup>

- $\|f\| > 0$  for  $f \neq 0$
- $\|0\| = 0$
- $\|a \cdot f\| = |a| \cdot \|f\|$  for all  $a \in \mathbb{C}^1$
- $\|f + g\| \leq \|f\| + \|g\|$  The Triangle Inequality.

For  $0 < p \leq \infty$  we will denote the set of all  $\mu$ -measurable functions,  $f$ , such that  $\|f\|^p$  is integrable by  $\mathcal{L}^p(X, \mathcal{F}, \mu)$ . We define a real-valued function,  $\|\cdot\|_p$ , on  $\mathcal{L}^p$  by ,

$$\|f\|_p = \left( \int \|f(x)\|^p d\mu(x) \right)^{1/p} .$$

For  $1 \leq p < \infty$  this function is called the  $\mathcal{L}^p$ -norm. For  $0 < p < 1$ ,  $\mathcal{L}^p$  is still a vector space, but  $\|\cdot\|_p$  is no longer a norm. Although  $\|\cdot\|_p$  can be used to define an interesting metric on  $\mathcal{L}^p$ , we will not consider this case further.

**Lemma 5.1.** *If the points are not collinear, then  $f$  is strictly convex.*

*Proof.* For  $X, Y, Q \in \mathbb{R}^n$ , and  $0 < t < 1$ , the Minkowski inequality shows,

$$\|X + Y\|_p \leq \|X\|_p + \|Y\|_p$$

implies,

$$\begin{aligned} & \|t(X-Q) + (1-t)(Y-Q)\|^p \\ & \leq \|t(X-Q)\|^p + 2\|t(X-Q)\| \cdot \|(1-t)(Y-Q)\| + \|(1-t)(Y-Q)\|^p \\ & = (t\|X-Q\| + (1-t)\|Y-Q\|)^p \end{aligned}$$

Hold if and only if  $X, Y, Q$  are not collinear, therefore,

$$f(tX + (1-t)Y) < tf(X) + (1-t)f(Y)$$

□

Since  $f(x_1, x_2, \dots, x_n)$  is convex, it proves that the Fermat point exists. Thus, by gradient descent

$$\nabla f = \frac{\partial f}{\partial x_l}(x_1, x_2, \dots, x_n = 0, l = 1, 2, \dots, n,$$

the function gets the minimum. It gives,

$$\frac{\partial f}{\partial x_l}(x_1, x_2, \dots, x_n) = \sum_{i=1}^m \sum_{j=1}^n k_i |y_l^i - x_l|^{p-1} \left( |m_j^i - x_j| \right)^{1/p-1} \frac{\sqrt{y_l^i - x_l}}{y_l^i - x_l} = 0, l = 1, 2, \dots, n.$$

Therefore, we have our generalized iterative algorithm,

$$T(x_l) = \frac{\sum_{i=1}^m \sum_{j=1}^n k_i x_l^i |y_l^i - x_l|^{p-1} \left( |y_l^i - x_j| \right)^{p/2-1}}{\sum_{i=1}^m \sum_{j=1}^n k_i |y_l^i - x_l|^{p-1} \left( |y_l^i - x_j| \right)^{p/2-1}}, \quad l = 1, 2, \dots, n.$$

## 6 The Implement Test

We build the PASCAL-based programs for iterative algorithms we proposed in the above sections. The following tables below are the test results for random test points. In norm-1 or  $\infty$ , the solution gives the exact result or range; in norm-2 and  $p$ , the programs also add the precision as a condition. We set the first iteration as the gravity point of the system.

<b>Input:</b>	Number of Points	Dimension	Weights
	4	6	2, 3, 1, 2
<b>Test Points:</b>	(4, 11, 3, 4, 5, 6), (13, 13, 2, 1, 5, 7), (17, 6, 8, 6, 7, 8), (8, 6, 4, 6, 7, 9)		
<b>Output:</b>	(8~13, 11, 3, 4, 5, 7)		

Table 1:  $\mathcal{L}^p$  Algorithm

<b>Input:</b>	Number of Points	Precision	Dimension	Weights
	6	0.001	2	1, 2, 4, 7, 6, 5
<b>Test Points:</b>	(4.71, -1.84), (-3.15, -2.44), (0.17, 2.99), (6.35, 2.86), (5.55, 2.44), (3.22, -2.56)			
<b>Output:</b>	Start : (1.480, -0.140)			
	1 : (1.713, 0.086)			
	2 : (1.763, 0.124)			
	3 : (1.776, 0.126)			
	4 : (1.782, 0.122)			
	5 : (1.787, 0.118)			
	6 : (1.790, 0.115)			
	7 : (1.793, 0.112)			
	8 : (1.795, 0.110)			
	9 : (1.798, 0.108)			
	10: (1.799, 0.106)			
	11: (1.801, 0.105)			
	12: (1.802, 0.104)			

Table 2:  $\mathcal{L}^2$  Algorithm

<b>Input:</b>	Number of Points	Norm	Precision	Dimension	Weights
	5	2.4	0.00001	5	5, 9, 1, 8, 6
<b>Test Points:</b>	(8, 5, 4, 8, 3), (3, 3, 7, 6, 3), (8, 7, 2, 6, 6), (4, 9, 3, 6, 2), (5, 6, 4, 5, 4)				
<b>Output:</b>	Start : (4.72414, 5.75862, 4.58621, 6.13793, 3.03448)				
	1 : (4.95849, 5.91367, 4.57670, 5.74708, 3.46078)				
	2 : (4.93330, 5.93940, 4.40332, 5.61458, 3.46066)				
	3 : (4.94537, 5.99067, 4.38619, 5.56072, 3.53263)				
	4 : (4.94590, 5.99950, 4.35138, 5.52565, 3.54720)				
	5 : (4.94774, 6.01196, 4.34204, 5.50728, 3.56604)				
	6 : (4.94833, 6.01319, 4.33222, 5.49539, 3.57347)				
	7 : (4.94886, 6.01513, 4.32840, 5.48840, 3.57989)				
	8 : (4.94909, 6.01562, 4.32523, 5.48396, 3.58315)				
	9 : (4.94928, 6.01613, 4.32359, 5.48123, 3.58549)				
	10: (4.94938, 6.01634, 4.32243, 5.47950, 3.58684)				
	11: (4.94945, 6.01651, 4.32176, 5.47842, 3.58773)				
	12: (4.94949, 6.01660, 4.32131, 5.47774, 3.58828)				
	13: (4.94952, 6.01666, 4.32104, 5.47730, 3.58864)				
	14: (4.94953, 6.01669, 4.32086, 5.47703, 3.58886)				
	15: (4.94954, 6.01672, 4.32075, 5.47686, 3.58900)				
	16: (4.94955, 6.01673, 4.32068, 5.47675, 3.58909)				
	17: (4.94956, 6.01674, 4.32063, 5.47668, 3.58914)				
	18: (4.94956, 6.01675, 4.32061, 5.47664, 3.58918)				
	19: (4.94956, 6.01675, 4.32059, 5.47660, 3.58920)				
	20: (4.94956, 6.01675, 4.32058, 5.47659, 3.58921)				

Table 3:  $\mathcal{L}^p$  Algorithm



## Acknowledgements

I would like to express my great gratitude to my adviser *Donglei Gu* and my friend *Wilhelm Braum*, for giving many insightful advices in this paper. Furthermore, I would like to thank my friend *Liaoyuan Shen* for the technical help in programming.

## References

- [1] Simpson, T. *The Doctrine and application fluxions*. Universidad de Salamanca, 1994.
- [2] Eiselt, Horst A., and Carl-Louis Sandblom. *Operations research: a model-based approach*. Springer, 2012.
- [3] Weiszfeld, Endre. "Sur le point pour lequel la somme des distances de n points donnés est minimum." *Tohoku Math. J* 43.355-386 (1937): 2.
- [4] Kuhn, Harold W. "A note on Fermat's problem." *Mathematical programming* 4.1 (1973): 98-107.
- [5] Chandrasekaran, R., and A. Tamir. "Open questions concerning Weiszfeld's algorithm for the Fermat-Weber location problem." *Mathematical Programming* 44.1-3 (1989): 293-295.
- [6] Vardi, Yehuda, and Cun-Hui Zhang. "A modified Weiszfeld algorithm for the Fermat-Weber location problem." *Mathematical Programming* 90.3 (2001): 559-566.
- [7] D. Rautenbach, M. Struzyna, C. Szegedy, J.Vygen, Weiszfeld's Algorithm Revisited Once Again, *Report No. 04946-OR*, Research Institute for Discrete Mathematics, University of Bonn, 2004.
- [8] Prugovecki, Eduard. *Quantum mechanics in Hilbert space*. Vol. 41. Access Online via Elsevier, 1982.