On minimizing distance by the road less traveled

Gregory P. Tollisen and Tamás Lengyel

Abstract

Optimization problems hold an extensive and honored position in creative problem solving. Often they can be reconceptualized so that the goal is to minimize the distances between mathematical objects of certain types. Frequently, for such problems, the standard techniques of calculus cannot be efficiently applied. Other optimization approaches, such as linear programming, may work, but without delivering insight into the problem or its solution. For the creative problem solver, these problems call for solutions that are lesser known and more unorthodox, even to the extreme of stepping outside of mathematics and allowing mechanical processes to provide the solution.

We look at four classes of problems each, more or less, resistant to solution by routine techniques of calculus alone. In Sections 1 and 2 we search for a best representative point for a point set. Finding a best fitting line for a point set is the topic of Section 3. In Section 4 we identify the integer lattice points closest to the origin by the l_1-norm, when constrained to a line. In each case, a sum of distances is to be minimized. Many examples of independent interest and broad appeal are also provided.

Our techniques arsenal draws from geometrical reflections and symmetries, repeated use of convex piecewise linear functions, Diophantine equations, the Euclidean algorithm and mechanical processes.

1 Finding an optimal location or representation

Finding optimal locations has been a popular subject in mathematics. There is extensive literature on the subject, in fact, over 3000 references are listed in [5]. We start with the one-dimensional version of the least absolute difference problem. Given a data set of real numbers, we want to find a value to represent the whole set. We consider two variants: one with equal and the other with given arbitrary positive weights.

Problem 1 *Let $x_1, x_2, \ldots, x_n$ be n real numbers. Find an a such that*

$$\min_a \sum_{i=1}^n |x_i - a| \tag{1}$$

is achieved.

All statistics students should know that the answer is the *median* of the values $x_1, x_2, \ldots, x_n$. When ordered from smallest to largest the median is defined to be

the middle point if n is odd and the arithmetic mean of the two middle points if n is even. That is not to imply that the median is the only solution, for when n is even, as we will see later, any value between the two middle ones is a solution.

A calculus based proof was presented by Cramér for any one-dimensional distribution in 1946. A pedagogically more instructive and simpler approach (e.g., [6]) can be given in the case of a distribution with probability mass located at a finite number of values. As Problem 1 is a special case of Problem 2, we will present only the solution to the latter one.

Problem 2 *Let $x_1 < x_2 < \ldots < x_n$ be n real numbers. Given the positive weights $w_1, w_2, \ldots, w_n$ find an a such that the minimum of*

$$D(a) = D(a; w_1, w_2, \ldots, w_n) = \sum_{i=1}^{n} w_i |x_i - a| \qquad (2)$$

is achieved.

Note that if $x_i = x_{i+1}$ then we can simply join these points into one with combined weight $w_i + w_{i+1}$.

Remark 1 The function $D(a)$ is non-negative, continuous and piecewise linear, so its minimum is attained at one of the points where the linear segments are joined, i.e., at some x_m.

Solution: The problem can easily be reduced to another one: find the minimum m such that $w'_m = \sum_{i=1}^{m} w_i / \sum_{i=1}^{n} w_i \geq 0.5$. The optimum value $a = x_m$ is sometimes referred to as the *weighted median* of the values x_i with weights w_i (apparently so named by Edgeworth in 1887 [2] in a slightly different context). To see this we need only to check the changes in the $D(a)$ as a moves from the left of x_1 to the right of x_n. In fact, a change from $x = a$ to $x = a+h, h > 0$, within the interval $[x_m, x_{m+1})$ results in the change $D(a+h) - D(a) = h(\sum_{i=1}^{m} w_i - \sum_{i=m+1}^{n} w_i)$, because x moves away from the first m points and towards the others. When m is increased by 1, one term is transferred from the second sum to the first, increasing the proportionality between the changes. Observing that the continuous $D(a)$ decreases provided that $D(a+h) - D(a) < 0$, we conclude the proof by increasing a until $D(a+h) - D(a)$ ceases to be negative.

Of course, a more calculus minded person can take the derivative of the continuous function $D(a)$ at all points except for critical points $x_i, 1 \leq i \leq n$. At the points where it exists, the derivative eventually changes from negative to positive, and therefore, a point a minimizing $D(a)$ can easily be found.   ■

The following remark provides us with a simple answer to many problems involving criterion (2).

Remark 2 If there exists an m such that $w'_m = 0.5$ then any point within the interval $[x_m, x_{m+1}]$ can be chosen for a. Otherwise, we have to select $a = x_m$. If there is a dominant weight with $w_m / \sum_{i=0}^{n} w_i > 0.5$ then $a = x_m$.

We will see that Examples 1, 2, 3, and 6 can be reduced to Problem 2, i.e., the weighted one-dimensional least absolute difference problem. The following example appears in [6].

Example 1 If asked where they would stand and wait for the next of three elevators unequally spaced along the wall, many students would choose to stand at the mean position. They do not recognize that standing at the mean minimizes the average squared distance and that the minimal average distance to the elevator is in fact achieved by standing at the median. If one's job is to carry heavy objects to elevators, one cannot ignore the distance from the initial point of arrival into the elevator area to the optimal place to wait. In our one-dimensional model, the point of arrival can either be to the left or right of, or somewhere in between the outermost elevators. Where should one stand in order to minimize the total distance, which we now redefine to include the additional distance from the point of entry to the point where one will wait?

"Surprisingly perhaps," as noted in [6], "the answer is to remain at the point of entry and to not move at all! ... one should stay put! ... In the case of an entry point that is to the left of the leftmost elevator, waiting anywhere between it and this elevator does not add unnecessary travel. A symmetric argument applies to entering the right side of the room."

The reason for the answer lies within Remark 2. In the general case of n elevators, the n trips are composed of $2n$ distances to be traveled: one each between the entry and the waiting point and from the waiting point to each elevator. The entry point comes with a weight of n.

Later, we will return to another nice example (Example 5) extending the previous one and showing a transition from the (weighted) median to mean.

Remark 3 In the case of equal weights we get $w_i' = i/n$, thus a can be selected to be the median of the data set. The previous remark on the possibility of several optimum positions applies and therefore, we can pick any value between the two middle ones if n is even.

We can view any value a for which the minimum is achieved in (1) or (2) as an optimal representation of the data set (with or without weights) with respect to the l_1-norm. This optimization problem can be generalized to $\mathbb{R}^2$.

The following example [15, Problem #469 on Siting a Central Depot] appears to be two-dimensional but in fact, it can be viewed as a nice combination of two instances of Problem 1 or, in the weighted version, Problem 2.

Example 2 The street plan of a city consists only of straight streets intersecting at right-angles, and at an odd number of the junctions there are kiosks. Figure 1 gives, as an example, a plan with ten streets and three kiosks. The occupants of the kiosks now wish to draw their wares from a common central depot. How should this be located so as to give a minimum total length for a single trip to the depot from each individual kiosk. The breadths of the streets may be neglected.

Figure 1

The solution can be easily described by a vector whose coordinates are formed by the medians in the corresponding directions. The solution can be generalized to the weighted version with an arbitrary finite number of points by taking the weighted medians. The same approach also works in higher dimension, for the l_1 or "taxicab" distances can be separated into one-dimensional absolute differences.

2 The two-dimensional version of the weighted least absolute difference problem

When the l_1 distance in Example 2 is replaced by Euclidean distance, we have the following restatement of Problem 2 in two dimensions.

Problem 3 *We replace x_i and a in Problem 2 by the plane vectors $\underline{v}_i = (x_i, y_i)$ and $\underline{a} \in \mathbb{R}^2$, respectively. Find an $\underline{a}$ such that the minimum of*

$$D(\underline{a}) = D(\underline{a}; w_1, w_2, \ldots, w_n) = \sum_{i=1}^{n} w_i |\underline{v}_i - \underline{a}| \tag{3}$$

is achieved.

This problem appears in the works of Steiner, Fermat, Weber, Torricelli, and Cavalieri. The optimal point is variously named after one or more of the mentioned people while the problem is often referred to as the Fermat–Weber location problem.

In passing to the next higher dimension, we generally lose the ability to reduce the problem to its one-dimensional form. (It is the sum of the *squared* Euclidean distances that can be separated into two one-dimensional problems, and is minimized by the center of mass. But see Example 3 below for an interesting and instructive exception.)

Furthermore, it is noteworthy that, unlike the one-dimensional version in Problem 2, the solution to Problem 3 is always unique, provided that the problem is indeed two-dimensional, in the sense that the points $\underline{v}_i$ are not collinear. For, if there were two distinct solutions, say $\underline{a}_1$ and $\underline{a}_2$, then $D(\underline{a})$ would be constant on the line segment $\overline{\underline{a}_1 \underline{a}_2}$ joining the solutions and thus $D(\underline{a})$ would have a directional second derivative equal to zero in either direction along the interior of this segment. This is because $D(\underline{a})$ is a convex function, since it is the sum of the convex functions $w_i |\underline{v}_i - \underline{a}|$. Now for each i, by virtue of its convexity, the directional second derivative of $w_i |\underline{v}_i - \underline{a}|$ is nonnegative in any direction at each point where it is

defined, and equals zero precisely when evaluated at a point on and in a direction along any line passing through $\underline{v}_i$. (Visually, the intersection of any vertical plane and the cone $w_i|\underline{v}_i - \underline{a}|$ is the upper half of a hyperbola, except when the plane passes through the axis of the cone.) Thus, for the directional second derivative of $D(\underline{a})$ to equal zero, the directional second derivative of each summand $w_i|\underline{v}_i - \underline{a}|$ must equal zero, and thus all of the points must lie on the line $\overline{\underline{a}_1\underline{a}_2}$.

No general technique is known to deliver a solution to this problem in closed form. Typically, the unique point that minimizes $D(\underline{a})$ is found numerically, for instance, by a method of deepest descent modified in order to account for the points where $D(\underline{a})$ is not differentiable, i.e., the points $\underline{v}_i$. (If in the process of deepest descent, we land on such a $\underline{v}_j$ for some j, we calculate the gradient $\nabla D^*(\underline{a})$ of the modified function $D^*(\underline{a}) = \sum_{i \neq j} w_i|\underline{v}_i - \underline{a}|$. If $|\nabla D^*(\underline{v}_j)| < w_j$ then the minimum occurs at $\underline{v}_j$. Otherwise, we continue in the direction of $\nabla D^*(\underline{v}_j)$.) Nature can perform the process of steepest descent as well, if the reader is willing to idealize nature. From an initial location, we allow a block to slide down the surface graph $(\underline{x}, D(\underline{x}))$, while applying just enough friction to the block so that its speed is infinitesimal throughout the process. The infinitesimal speed removes kinetic energy, and thus, the block slides in the direction that most rapidly decreases potential energy.

The Fermat–Weber problem can be solved similarly in higher dimensions as well by deepest descent techniques such as the popular Weiszfeld (Vázsonyi) iterative method and its modifications.

We present two examples of Problem 3 with two different outcomes. The first reduces to the one-dimensional case but the second does not. However, a *mechanical analysis* that we refer to as the *"Steinhaus method"* can be applied to both.

Example 3 [14, p. 172] The U.S. Chess Federation is planning a tournament for all the Grandmasters in the country. It wants to locate the tournament so the total distance traveled by all players is a minimum. More than half of the Grandmasters live in New York City. Can you *prove* that the best site to hold the tournament is in New York City regardless of the locations of the other players? (Assume that New York City is a point, that the U.S. is flat, and the players always take direct routes and never get caught in New York traffic.)

The answer is yes. Before explaining this somewhat surprising result and why it reduces to the one-dimensional case, we include as our second example a vintage problem where no position is assigned a majority.

Example 4 [12, pp. 113-6] Three villages are to build a common school. In order to reduce as far as possible the total time spent by pupils in traveling to school, they have to find an appropriate spot for the location of the school. They have, for instance, 50, 70, and 90 children, respectively.

We illustrate the Steinhaus method by the following

Solution to Example 4 [12]: "Stretching out the map of the district on a table, we make holes in the table where the villages are, pass three strings through the holes, tie the upper ends into a knot, and weight the lower ones with 50, 70, and 90 ounces, respectively. The school should be built where the knot is caught."

The above solution is also based on the concept of potential energy. The physical system is in equilibrium if its potential energy is minimum. Since the total length of the strings is a given constant, this minimum is achieved if a minimum length of the strings is horizontal in the case of equal weights as in the previous example. A similar argument applies to the situation of unequal weights.

While not intending to discourage the enthusiastic hole driller, this technique clearly works for an arbitrary number of holes. The above physics inspired idea also can be applied to solve other geometric problems (cf. [12] and [13]).

Now we return to the *chess* problem and readily obtain a

Solution to Example 3 by the Steinhaus method: It is easy to see that if one village has more than 50% of the students then this solution will move the knot to that village. The same applies to the best location for the tournament with the majority of players living in one city. ∎

We can also find an analytical solution to the *chess* problem using Problem 2.

Alternative **Solution** to Example 3: Let $\underline{v}_k$ be the location with the majority of players. Assume that the best tournament location is at $\underline{a} \neq \underline{v}_k$. We take the line l through $\underline{a}$ and $\underline{v}_k$. Now we project every point onto line l. The one-dimensional solution (cf. Problem 2 and Remark 2) guarantees that we can improve $D(\underline{a})$ by moving $\underline{a}$ closer to $\underline{v}_k$ on line l. ∎

In his book [12], Steinhaus translates the mechanical method to its vector form for 3 points and arbitrary weights. It is based on the notion of *reciprocal figures* of Cremona (or *scheme of forces*) from the theory of graphical statics. An auxiliary triangle is constructed with sides 50, 70, and 90. The claim is that the school should be built at the point from which the villages are visible in directions that form the same angles as the exterior angles of the auxiliary triangle. Despite its beauty, this method has a limited scope as it often fails to solve the problem; for example, when the sides do not satisfy the triangle inequality (which happens precisely when there is a majority).

It is worth making a few comments on the version with equal weights (equivalent to the unweighted version). For equal weights $w_i = w, i = 1, 2, \ldots, n$, with $n = 3$ and 4 we could use purely geometrical arguments without auxiliary triangles.

For example, if the $n = 3$ points form an acute triangle then the auxiliary triangle gives the optimal solution at point P with the property that each side spans a 120° angle when viewed from P. However, this can also be demonstrated by a clever proof involving rotation. Assume that the locations are represented by the vertices A, B, and C of the $\triangle ABC$ (labeled in the clockwise direction) and that the optimal point is P. We rotate the $\triangle APC$ around point A through 60° in

the clockwise direction to obtain the $\triangle AP'C'$. It is easy to see that both P and P' must be points on the line $C'B$. It follows that $\angle APB = 120°$ and similarly, $\angle APC = \angle BPC = 120°$. For obtuse triangles the optimal point is found at the vertex with the obtuse angle.

The solution is even simpler in the case of $n = 4$. The triangle inequality, applied to triangles formed by the diagonally opposite vertices and the solution point, shows that for a convex set the optimal solution is given by the intersection of the diagonals. By a similar argument, for a concave quadrilateral the point with the concave angle provides the optimum point.

The location problem (Problem 3) asks for a network of minimum total length connecting all points with the help of a single extra point. It can be generalized to the so called Steiner tree problem in which we allow the addition of an arbitrary number of points rather than just one to find the shortest network connecting all points. Conveniently, the Steiner problem can also be solved by nature. Place upright pegs representing the points between two transparent parallel plates. Dip the model into a soap solution and remove. The surface tension of the soap film formed settles at the stable equilibrium by minimizing the area of the film. Since the distance between the surfaces is constant, the film will also minimize the total distance, forming Steiner points in the process.

Now we return to a variation of Example 1 in which Steinhaus' approach is used to find the optimum solution. Furthermore, we derive some intriguing aspects of the optimum solution by analyzing the forces involved at equilibrium.

Example 5 Suppose we modify Example 1 by placing all three elevators along one wall of a corridor and requiring our passenger to wait along the opposite wall. We assign the weight w_i to the ith elevator to reflect, say, how frequently it is being used. For instance, in Example 1 the elevators have weight 1 and the entry point has weight 3. How do we locate the optimum position for the passenger?

We prepare to apply the mechanical approach by reflecting the three elevators across the opposite wall, and dividing the weight of each elevator equally with its own reflection, increasing the total number of locations to six. If one applies the Steinhaus method to the six locations, the knot will lie along the opposite wall by symmetry, and the total weighted distance will be minimum and equal to the total weighted distance for the original problem.

We continue for any number of elevators. The components of the vector forces applied at the knot parallel to the corridor walls must equal zero at equilibrium: $0 = \sum w_i \sin \theta_i$, where θ_i is the signed angle according to the righthand rule made by the vector force corresponding to the ith elevator with respect to a perpendicular to the walls. The result is a (sort of) Snell's law. (In fact, Snell's law can be derived by defining, as usual, the path taken by a light ray through the interface of two transparent media as the path of least time. Then, the path taken by the light between two points separated by the interface is just the path of minimum

weighted distance from a point on the interface to the two points, where the weight for each point is the reciprocal of the velocity of light in the medium occupied by the point. Reflecting one of the points across the interface reduces the situation to our modified elevator problem with two elevators.)

The limiting cases for the width l of the corridor are particularly interesting. As $l \to \infty$, $h_i \sim l$ where h_i is the distance between the ith elevator at location x_i on the wall and the optimum point on the opposite wall. Then,

$$0 = l \cdot \sum w_i \sin \theta_i = -l \cdot \sum w_i \frac{x_i - a}{h_i} = -\sum \frac{l}{h_i} \cdot w_i(x_i - a) \sim -\sum w_i(x_i - a)$$

where a is the location of the optimum point projected onto the wall occupied by the elevators. Thus in this case, the limiting value of a is the weighted average of the locations of the elevators.

When $l = 0$, the problem reduces to Problem 2 with the optimal position being the median. As highlighted in Remark 2, when $w'_m = 0.5$, the optimum position is not unique, and in fact, any position between x_m and x_{m+1} can serve equally well. However, by using a second order argument, we can see that allowing l to approach zero selects a unique optimum position with a curious physical interpretation. We let $\phi_i = \frac{\pi}{2} - |\theta_i|$ and observe that, $\cos \phi_i - 1 \sim -\frac{\phi_i^2}{2} \sim -\frac{l^2}{2(x_i-a)^2}$ as $l \to 0$. Then a rewriting of Snell's law gives us

$$0 = \frac{1}{l^2} \cdot \sum \text{sign}(x_i - a)w_i \cos \phi_i$$

$$= \frac{1}{l^2} \cdot \sum \text{sign}(x_i - a)w_i + \frac{1}{l^2} \cdot \sum \text{sign}(x_i - a)w_i(\cos \phi_i - 1)$$

$$\sim 0 + \frac{1}{l^2} \cdot \sum \text{sign}(x_i - a)w_i \left(-\frac{l^2}{2(x_i - a)^2} \right) \sim -\frac{1}{2} \sum w_i \frac{\text{sign}(x_i - a)}{(x_i - a)^2}.$$

Thus, in the case $w'_m = 0.5$, if a positive charge of measure w_i were placed at location x_i for each i, then the limiting position a that minimizes the total weighted distance is the stable equilibrium position of a positive charge between positions x_m and x_{m+1}. ∎

Note that if we replace the l_2-norm (in the above setting) by the l_1-norm then the optimum is always on the far side across from the weighted median.

We consider another problem which, though dealing with a best fitting line, a topic covered in Section 3, can also be reduced to Problem 2. This example is a restricted version of Problem 5 of Section 3, as we look for a least absolute deviation line which goes through the origin.

Example 6 For given points $(x_i, y_i) \in \mathbb{R}^2$, $1 \le i \le n$, find $\min_m \sum_{i=1}^n |m x_i - y_i|$.

Note that any point with $x_i = 0$ contributes a fix $|y_i|$ to this sum, so it can be ignored. Therefore, we consider $\min_m \sum_{i=1}^n |x_i||m - y_i/x_i|$ with $x_i \ne 0, 1 \le$

$i \leq n$. This problem is a special case of the weighted one-dimensional problem (Problem 2) with weights $|x_i|$ and data points $y_i/x_i, i = 1, 2, \ldots n$. Therefore, the optimum is achieved when the best fitting line goes through some point (x_k, y_k), i.e., $m = y_k/x_k$ for some $k : 1 \leq k \leq n$. It is not true in general that about 50% of the points are below and above the line.

This problem can be generalized by translating the origin into any given pivot point. If the given pivot point is the data point (x_j, y_j) then we need the weighted median of the slope values $\frac{y_i - y_j}{x_i - x_j}, i = 1, 2, \ldots n$, but $i \neq j$, with weights $|x_i - x_j|$. A setting where this problem might arise is when one wishes to find the best fitting line in the least absolute deviation sense, given that one of the data points is known with certainty to be on the line.

3 Finding the best fitting line: the least absolute deviation line

Fitting a straight line to a finite collection of data points in the plane is a fundamental problem in statistical estimation. Although methods such as least squares are well understood and easy to compute, these methods are known to suffer from the phenomenon that a small number of outlying points can perturb the line of fit by an arbitrarily large amount [10].

A popular alternative is the least absolute deviation (LAD) line which is less sensitive to the presence of outliers (see Problems 4 and 5). For historical background of this and other best fitting line criteria and techniques the reader can consult [2]. It is worth noting that the problem of finding the LAD line (in the simplified form mentioned in Example 6) is older than that of the least squares.

From a technical point of view LAD line fitting embodies a set of ideas that is important in linear optimization theory and numerical analysis [2]. Several refined and sophisticated algorithms have been developed to find an LAD line for a given set of data points. We leave to other sources the task of evaluating and comparing these algorithms [1, 2]. Our intent is the extraction of enough properties to buy us an easy characterization of all LAD lines for a given data set, and to suggest two simple algorithms for finding them that flow out of our analysis.

Note that when there are several LAD lines then we can take their averages to yield other LAD lines [1]. In fact, as will be seen in this section, any convex combination will also result in an LAD line.

The next problem involves three points only but it illuminates the main idea.

Problem 4 [11, Problem E3079] *Let (x_1, y_1), (x_2, y_2) and (x_3, y_3) be three points in $\mathbb{R}^2$, no two of which have the same x-coordinate. Find a least absolute deviation (LAD) line $y = Ax + B$, i.e., find $(a, b) = (A, B)$ that minimizes*

$$g(a, b) = \sum_{i=1}^{3} |ax_i + b - y_i|.$$

Must an LAD line pass through two of the three points?

Solution: The answer is yes. Here is a sketch of the proof. A line that does not contain any of the three data points, when shifted vertically toward the majority of the points, will decrease $g(a, b)$. Thus, one of the data points must lie on the LAD line. If the line does not contain another point, then it is easy to see that we can further decrease $g(a, b)$ by rotating the line around this pivot point until we meet another point. Of course, Example 6 also guarantees that the line contains a second data point, too. In fact, a simple geometric argument shows that the LAD line must go through the two points with the left- and rightmost x-coordinates. ∎

When the condition that no two of the three data points have the same x-coordinate is replaced by the weaker condition that all three data points must not share the same x-coordinate, it is easy to find examples of Problem 4 where the LAD line is not unique. Nevertheless, the analysis illustrated above guarantees that at least one LAD line must pass through two data points. If all x-coordinates are the same then this problem reduces to Problem 1 with a in (1) playing the role of the intercept at the common x-coordinate.

The following is a generalization of Problem 4 to n points under the weaker condition.

Problem 5 *Let (x_1, y_1), (x_2, y_2), ..., (x_n, y_n) be n points in $\mathbb{R}^2$, not all sharing the same x-coordinate. Find a least absolute deviation (LAD) line $y = Ax + B$ for the set of points, i.e., find $(a, b) = (A, B)$ that minimizes*

$$g(a, b) = \sum_{i=1}^{n} |ax_i + b - y_i|.$$

We first note that prior knowledge of the optimum intercept b or slope a reduces the search for an LAD line to Problem 2 (in the former case via Example 6).

The graphical argument in Problem 4 (taking into account the weakened hypothesis) can be followed here without difficulty to conclude that an LAD line exists that passes through two data points. We will call any such LAD line a *special LAD line.* Alternatively, one could come to the same conclusion by noting that $g(a, b)$ is piecewise linear in either variable when holding the other fixed and thus attains its minimum at the zero of one of its summands. The result immediately suggests

Algorithm 1 for finding an LAD line: Run through all pairs of points and check which of the determined lines attains the optimum absolute deviation.

More is to be gained. The function $g(a, b)$ is a convex function, so the set S on which it attains its minimum is convex. Any convex combination of LAD lines corresponds to a convex combination of points in S, and thus is itself an LAD line. But $g(a, b)$ is piecewise linear as well, in the sense that its domain can be

partitioned, by the n lines $b = -x_i a + y_i$ in the ab-plane, as it will be explained in the proof of Claim 2, into a finite number of convex polygonal regions (some of which are infinite in extent), and within each region $g(a, b)$ is linear. Thus, the set S on which $g(a, b)$ attains its minimum consists of a single point, a line segment, or a convex polygonal region of finite extent (since $g(a, b) \to \infty$ as $|a| + |b| \to \infty$). In any case, the vertices (or single point) of the set S are among the points of intersection of the n lines, and the coordinates of each give the slope and intercept of a special LAD line in the xy-plane, i.e., one passing through two data points. Thus, we can make

Claim 1 The set of all LAD lines is precisely the convex hull of the special LAD lines.

Since the special LAD lines are now seen to play a crucial role in characterizing all LAD lines, there arises the question of how many special LAD lines can be found for a given set of data points. Interestingly enough, the answer to this question pivots on whether the number of data points is even or odd. In the even case, the possible number of special LAD lines is unbounded. In fact, for any given even number $n = 2k$, a data set can be constructed with exactly n special LAD lines. For example, we can select k points along the convex upper branch of the hyperbola $y^2 - x^2 = 1$, and their k reflections with respect to the x-axis along the concave lower branch. Then, the special LAD lines are precisely the $k - 1$ lines passing through adjacent points of the convex branch, their reflections, the line passing through the left-most point of the convex branch and the right-most point of the concave branch, and its reflection.

In light of the preceding considerations, the answer to the odd case may appear surprising. Our conclusion will be

Claim 2 If the data set has an odd number of points, then there are at most two special LAD lines for the set. Moreover, if there are two special LAD lines, they must intersect, and every other LAD line must pass through the same intersection.

It is illuminating to reach our conclusion by employing more completely the dual relationship between the xy-plane and the ab-plane. First, note that every line $y = a_0 x + b_0$ in the xy-plane corresponds to a point (a_0, b_0) in the ab-plane. Further, every point (x_0, y_0) in the xy-plane corresponds (by way of the lines $y = ax + (y_0 - ax_0)$ passing through it) to the line $b = -x_0 a + y_0$ in the ab-plane with slope $-x_0$ and vertical intercept y_0. Now, we translate the absolute deviation function $g(a, b)$ directly into the ab-plane by noting that the absolute deviation for the particular line $y = a_0 x + b_0$ in the xy-plane is

$$g(a_0, b_0) = \sum_{i=1}^{n} |a_0 x_i + b_0 - y_i| = \sum_{i=1}^{n} |b_0 - (-x_i a_0 + y_i)|.$$

Interpreted in the ab-plane, the second sum is the sum of the vertical distances

between the point (a_0, b_0) and the n lines $b = -x_i a + y_i$, which we will call the *absolute deviation* between the point and the set of lines.

An example may be in order. A data set in the xy-plane is given in Figure 2. The points p through t correspond to the lines in the plane shown in Figure 3, also labeled p through t. The line $y = -x + 1$ that passes through points p and r in the xy-plane corresponds to the point $(-1, 1)$ in the ab-plane which, not surprisingly, is the intersection of the two lines labeled p and r. A quick calculation in each plane will show that the absolute deviation is 8.

We search for the *least* absolute deviation in the ab-plane. The absolute deviation for any point is improved by moving the point vertically to the majority of lines until it is on the "middle" line, with as many lines above as below. Thus, the least absolute deviation can be found along the piecewise linear *middle path*. (See the highlighted path in Figure 3). Define the function $h(a)$ to be the absolute deviation for the point on the middle path corresponding to a. Notice that, except for a finite number of values for a, the number of lines is the same above and below the point (a, b) on the middle path, and $h(a)$ can be more easily calculated. In fact, we can find $h(a)$ by taking the difference between the sum of the b-coordinates for the points on the lines above the middle path at a and the sum of the b-coordinates for the points on the lines below the middle path at a. The function $h(a)$ itself is a piecewise linear and convex function (and is illustrated in Figure 5 for the example). The piecewise linearity for $h(a)$ should be obvious by now. The convexity is established by showing that the derivative $h'(a)$ strictly increases across each value a_0 where the middle path is intersected with another line or lines. At such a point (a_0, b_0) in the ab-plane, the slope of any line not intersecting the path makes no contribution to the *change* in $h'(a)$. As for those lines passing through (a_0, b_0) the number of lines above the middle path remains the same to the left and right of a_0, as well as the number of lines below the middle path. Now, order the slopes of the lines passing through (a_0, b_0) from least to greatest: $m_0 < m_1 < \ldots < m_k$ with m_l and m_{k-l} being the slopes of the middle path entering and exiting the point (see Figure 4). The slopes' contributions to the change in $h'(a)$ come in pairs, working from outside towards the middle of the list. For $i < \min(l, k-l)$, the contribution of the ith and $(k-i)$th slopes to $h'(a)$ is $m_i - m_{k-i}$ upon entering (a_0, b_0) (because the line with slope m_i is above the middle path and the line with slope m_{k-i} is below it), and $m_{k-i} - m_i$ upon exiting it. The difference $2(m_{k-i} - m_i)$ is positive. For $\min(l, k-l) < i \leq \frac{1}{2}k$, the lines (or line when $i = \frac{1}{2}k$) with slopes m_i and m_{k-i} fall on the same side of the middle path and add no contribution to the change in $h'(a)$. Finally, the contribution of m_l and m_{k-l} to the change in $h'(a)$ is $|m_l - m_{k-l}|$. Since there are at least two lines passing through (a_0, b_0), the sum of the individual contributions must be positive. Thus, $h(a)$ is convex in such a way that slope of each linear section is strictly greater than the slope of the preceding section. The function $h(a)$ must therefore take on its minimum at strictly one point, corresponding to a special LAD line, or along one linear section, with only its two endpoints corresponding to special LAD lines. In either case there can be no more than two special LAD lines. ∎

It is not hard to see that for an even number of data points a middle path can be found in the ab-plane as well, though it is not uniquely determined, and along any one of these middle paths, an LAD line will be found. Thus, for any number of points, we have the following improvement over our previous algorithm.

Algorithm 2 for finding all special LAD lines: For each data point, calculate the slopes of the lines passing through this point and each of the other data points. Then, among these lines, calculate the absolute deviations only of the (at most two) lines with median slopes. The special LAD lines are precisely those calculated with minimum absolute deviations.

| Figure 2 | Figure 3 |

| Figure 4 | Figure 5 |

In higher dimensions the function to be minimized would be

$$g(c_0, c_1, \ldots, c_{k-1}) = \sum_{i=1}^{n} \Big| \sum_{j=1}^{k-1} c_j x_{i,j} + c_0 - y_i \Big| \tag{4}$$

in variables $c_0, c_1, \ldots, c_{k-1}$. Gauss formalized the corresponding statement for the higher dimensional case in 1809. As a matter of fact, for the data set $X = \{\underline{v}_i\}_{i=1}^n$, with $\underline{v}_i = (x_{i,1}, x_{i,2}, \ldots, x_{i,k-1}, y_i), i = 1, 2, \ldots, n, (X \subset \mathbb{R}^k, k \leq n)$, there is an optimum vector $\underline{c} = (c_0, c_1, \ldots, c_{k-1})$ with respect to function g defined in (4) so that at least k points are contained in the hyperplane $\sum_{j=1}^{k-1} c_j x_{i,j} + c_0 - y_i = 0$ [2, pp. 7-8].

4 A discrete optimization problem in two variables

In this last section, we will present a problem of a nature seemingly unrelated to those we have already considered, having to do with linear Diophantine equations. Indeed, in certain cases, the extended Euclidean algorithm immediately provides the solution. Yet, as will be seen, the algorithm must be complemented in the general case with the evaluation of a piecewise linear function as in Problems 1 and 2, but now at two consecutive integers.

Here we find a solution which is nearest to the origin with respect to the least absolute difference. The following problem can be found in various sources (e.g., `http://cut-the-knot.com/water2.html#proof`).

Example 7 For a perfect breakfast, a fellow decides to boil an egg exactly 15 minutes. He has two hour glasses–one for 7 minutes, another for 11. How should he go about preparing his breakfast? What would be the minimum required number of glass turns? (Starting does not count as "turning.")

It takes one turn of the larger glass. We start both glasses. After the smaller glass empties we start boiling our egg and continue watching the big glass. When the big glass runs out the egg will have been boiling for 4 minutes. We turn the big glass and wait another 11 minutes. We get $4 + 11 = 15$ with one turn of the big glass. Clearly, it is an absolute minimum.

In the generalized version, there are two hour glasses with capacities a and b $(b > a)$ minutes. It is possible to measure $2b - a$ minutes by the above mentioned steps with a and b playing the role of 11 and 7.

Bogomolny refers to [15, #479] and to Quicky #1 [3, pages 190 and 199] as the potential sources for this example. Note that in the second reference Gardner, who cites other sources, seems to have in mind a variation where the shortest time is to be found. However, because it fits the theme of our discussion, we will pursue only the version already proposed.

In our discussion, notice that both hour glasses were run concurrently and without interruption. We restrict the scope of possible solutions beyond that of the original problem by stipulating that

(1) each hour glass, if used, must be used for an uninterrupted period of time by immediately turning it over each time it runs out during that period, and that

(2) the two hour glasses either must be started simultaneously or one started immediately after the other has run out for the last time.

Under these restrictions, each solution is simply the solution to a linear Diophantine equation.

Example 8 Generate all integer solutions in x and y to the following problems

(a) $1 = 11x + 7y$

(b) $15 = 11x + 7y$,

and find x and y such that $|x| + |y|$ is minimum.

Part (b) corresponds to Example 7. For simplicity, we count the number of *runs* of each hourglass which corresponds to $|x| + |y| - 2$ turns in the sense of the above example, except if one hourglass is never run.

Solution: We use the parameterized form of the solution obtained by the *extended Euclidean algorithm*. For example, $\gcd(11, 7) = 1 = 11x + 7y$ with $(x, y) = (2, -3)$, and the general solution is $x = 2 + 7t$ and $y = -3 - 11t$ with an arbitrary integer

t. Note that the next to last convergent $3/2 = [1; 1, 1]$ in the continued fraction expansion of $11/7 = [1; 1, 1, 3]$ provides us with the initial solution $(2, -3)$ or $(-2, 3)$ yielding the pair $\pm \gcd(11, 7)$. Now we are ready to solve (a) and (b).

(a) We need that $\min |2 + 7t| + | - 3 - 11t| = 2 + 3 = 5$ meaning that $(x, y) = (2, -3)$, with $t = 0$, will provide the optimal solution.

(b) We can multiply the equality $1 = 11 \cdot 2 - 7 \cdot 3$ by 15 to get an initial solution $(30, -45)$, i.e., $15 = 11 \cdot 30 - 7 \cdot 45$. We need $\min |30 + 7t| + | - 45 - 11t|$ which is $|2| + | - 1| = 3$ and achieved at $t = -4$. It gives the optimal solution $15 = 11 \cdot 2 - 7 \cdot 1$ at $(x, y) = (2, -1)$ confirming the solution of Example 7.

Now we generalize Example 7 under assumptions (1) and (2).

Problem 6 *In general, find a solution to the Diophantine equation $k = ax + by$ ($k \geq 0$, a and b relatively prime positive integers), that minimizes $|x| + |y|$.*

Solution: Note that if $b = 1$ then the solution is trivial. By excluding the simple case where $a = b$, without loss of generality, we can assume that $2 \leq b < a$, and that by some method such as the one illustrated in Example 8, we have already found an initial solution (k_x, k_y) to the Diophantine equation. Then, we get all solutions parametrically:

$$x(t) = k_x + bt, \quad y(t) = k_y - at, \qquad (5)$$

as t runs through the integers.

Our goal is to minimize the function $f(t) = |x(t)| + |y(t)|$. The approach we take is analytically geometrical, viewing all of the solutions to the Diophantine equation as equally spaced lattice points in $\mathbb{R}^2$ on the line $y = -\frac{a}{b}x + \frac{k}{b}$, which has slope $-\frac{a}{b} < -1$ and passes through the first quadrant. Notice that, for $D > 0$, the level curve $|x| + |y| = D$ is a diamond oriented square centered at the origin with opposite vertices at $\pm D$ on the two axes. By increasing D from a small enough value, we expand the diamond until it first intersects the line at its x-intercept $\frac{k}{a}$, and then if necessary, continue to expand the diamond until a solution to the Diophantine equation (x^*, y^*) is reached for the first time (cf. Figure 6). The solution minimizes $f(t) = |x(t)| + |y(t)|$ and further, $|y^*| < a$, which suggests the following simple algorithm.

From an initial solution (k_x, k_y) to the Diophantine equation $k = ax + by$, subtract a multiple $t = m$ of a from k_y so that $0 \leq k_y - am < a$, thus minimizing $y(t) \geq 0$. Then check the solutions $(k_x + bm, k_y - am)$ and $(k_x + b(m + 1), k_y - a(m + 1))$ to determine which one minimizes $|x| + |y|$. Even more briefly, find $m = \lfloor \frac{k_y}{a} \rfloor$, and choose the solution corresponding to the minimum between $f(m)$ and $f(m + 1)$. ∎

We note that the steps in the Euclidean algorithm for finding $\gcd(a, b)$ are paralleled by calculating the partial quotients of the continued fraction expansion $[a_0; a_1, \ldots, a_n]$ of a/b. We can assume that $a_n \geq 2$. Note that the $r + 1$st ($r \geq -1$)

convergent $p_{r+1}/q_{r+1} = [a_0; a_1, \ldots, a_{r+1}]$ is generated from the preceding convergents by the matrix identity

$$\begin{pmatrix} p_{r+1} & p_r \\ q_{r+1} & q_r \end{pmatrix} = \begin{pmatrix} p_r & p_{r-1} \\ q_r & q_{r-1} \end{pmatrix} \begin{pmatrix} a_{r+1} & 1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} p_0 & p_{-1} & p_{-2} \\ q_0 & q_{-1} & q_{-2} \end{pmatrix} = \begin{pmatrix} a_0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \quad (6)$$

and that $\gcd(a, b) = |p_{n-1}b - q_{n-1}a|$.

Claim 3 In the special case that $k = \gcd(a, b)$, the unique optimal solution is $(q_{n-1}, -p_{n-1})$ if n is odd and $(-q_{n-1}, p_{n-1})$ if n is even.

To prove this claim we note that, as we mentioned above, the next to last continued fraction convergent $c/d = p_{n-1}/q_{n-1}$ of $a/b = p_n/q_n$ gives us an initial solution $(k_x, k_y) = (d, -c)$ or $(-d, c)$ to $k_x a + k_y b = k$. Since $p_r, r \geq 0$, and $q_r, r \geq 1$, are strictly increasing sequences by (6), we observe that $c < a$ and $d < b$. In fact, in this case the unique minimum is taken at (k_x, k_y) with value $f(0)$. For, in the case when $k_y = -c < 0$ we get $f(-1) = -d + b - c + a > f(0) = d + c$, for recurrence (6) yields $a > a_n c \geq 2c$ and $b \geq a_n d \geq 2d$ and thus $a + b > 2(c + d)$. Similarly, if $k_y = c > 0$ then $f(1) = -d + b - c + a > f(0) = d + c$ for the same reason. ■

The visually minded reader can use the Stern-Brocot tree [4, pages 119 and 305] to derive an alternative proof. Note that from the general solution (5) for $(x(t), y(t))$ we can form the linear rational transformation $g(t) = \frac{c+at}{d+bt} = [a_0; a_1, \ldots, a_n + \frac{1}{t}]$ with integer t. If $t = 0$ then we get $g(0) = c/d$, i.e., the initial solution (k_x, k_y) in the fractional form while $g(t)$ tends to a/b as $t \to \infty$ or $t \to -\infty$. In the tree, the general solution can be identified as the union of the two points corresponding to p_{n-1}/q_{n-1} and $(p_{n-1}(a_n - 1) + p_{n-2})/(q_{n-1}(a_n - 1) + q_{n-2})$, i.e., $t = 0$ and $t = -1$ in (5), respectively; and the two paths starting immediately below p_n/q_n and going toward the value of p_n/q_n. In fact, these branches correspond to $g(t) = [a_0; a_1, \ldots, a_n, t]$ for $t \geq 1$, and $g(t) = [a_0; a_1, \ldots, a_n - 1, 1, -(t+1)]$ for $t \leq -2$. The structure of the tree guarantees the claim.

We note that the values of $g(t)$ at fractions $t = -1, -1/2, -1/3, \ldots, -1/(a_n - 1)$ are called intermediate fractions [7], and they represent the numbers on the path between p_{n-1}/q_{n-1} and p_n/q_n.

The continued fraction context makes clear the otherwise remarkable fact that the same optimum solution is derived for any pair (a, b) where the partial quotients are the same except for the last one, provided $\gcd(a, b) = k$. Interested readers can consult [7] or [8] for a comprehensive study of continued fractions.

The next example shows that multiple optima might exist in Problem 6.

Example 9 Let us consider the Diophantine problem with $k = 9$ and $(a, b) = (11, 7)$. We get $1 = 2 \cdot 11 - 3 \cdot 7$, yielding $(18, -27)$ as an initial solution (k_x, k_y) and general solution $(x(t), y(t)) = (18 + 7t, -27 - 11t)$ to $9 = 11x + 7y$ after multiplication by k. For the "new" solution we get $(-3, 6)$ with $t = m = \lfloor \frac{-27}{11} \rfloor = -3$. Now $f(-3) = |-3| + 6 = 9 = 4 + |-5| = f(-2)$ yielding a multiple optimum at $t = -3$ and $t = -2$ (see Figure 7).

We can characterize the cases with multiple optima by

Claim 4 Let $C = (a+b)/2$ be. Multiple optima exist if and only if $x = \frac{k-bC}{2C}$ is a non-positive integer and then $D = C$, or $x = \frac{k-bC}{a}$ is a non-negative integer and then $D = x + C$.

Note that Levit studied a related problem in [9]. He calls a solution to $ax + by = k, |a| \geq |b|$, a minimum solution with respect to x if $|x|$ assumes its minimum, and a definitely least solution (*dls*) if $|x|$ and $|y|$ both attain their least possible values. An algorithm presented in [9] finds a *dls* if one exists with a simple test to check if it is a *dls* indeed (cf. Corollary 2 and Lemma in [9]). If a *dls* exists (e.g., when a is odd and $k < C$ [9, Theorem 3]) then it clearly minimizes $|x| + |y|$, but otherwise (e.g., when a is odd and $k > ab - C$, ibid.), this approach provides little help in solving Problem 6.

We note that Problem 6 is a special instance of the one-dimensional version of the celebrated *Closest Vector Problem*, for we try to find the closest vector of the integer lattice formed by $\{t(7, -11) | t \text{ integer}\}$ to $(2, -3)$.

Figure 6   Figure 7

---

5   Conclusion

Optimization questions can be found in many fields of mathematics. We have highlighted a selection of problems a little off of the beaten path, taken from statistics and recreational mathematics, that are interesting in their own right and worthy of greater exposure. In addition, we hope to have introduced to the neophyte and recalled to the experienced problem solver alternative techniques that suggest the rich variety of approaches for optimization, even to the crossing over of disciplines.

6   Acknowledgment

References

1. David Birkes and Yadolah Dodge, *Alternative Methods of Regression*, Wiley, New York, 1993.

2. Peter Bloomsfield and William L. Steiger, *Least Absolute Deviations, Theory, Applications and Algorithms*, Birkhäuser, Boston, 1983.

3. Martin Gardner, *Mathematical Circus*, MAA, 1992 (updated and revised from the 1981 edition published by Vintage Books, NY).

4. R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics*, 2nd edition, Reading, MA: Addison-Wesley, 1994.

5. T. Hale, Trevor Hale's location science references,

   `http://www.ent.ohiou.edu/~thale/thlocation.html`

6. James A. Hanley, Lawrence Joseph, Robert W. Platt, Moo K. Chung, and Patrick Bélisle, Visualizing the Median as the Minimum-Deviation Location, *Amer. Stat.* 55(2001), 150–152.

7. A. Ya. Khinchin, *Continued fractions*, New York, Dover, 1997.

8. D. E. Knuth, *The Art of Computer Programming*, vol. 2, Seminumerical Algorithms, 2nd edition, Reading, MA: Addison-Wesley, 1981.

9. R. J. Levit, A minimum solution of a Diophantine equation, *Amer. Math. Monthly* 63(1956), 646–651.

10. Jiří Matoušek, David Mount, and Nathan S. Netanyahu, Efficient randomized algorithms for the repeated median line estimator, *Algorithmica* 20(1998), 136–150.

11. Solution to Problem E3079, *Amer. Math. Monthly* 94(1987), 1000–1002

12. Hugo Steinhaus, *Mathematical Snapshots*, Dover, 1999.

13. Tadashi F. Tokieda, Mechanical ideas in geometry. *Amer. Math. Monthly* 105(1998), 697–703.

14. Ravi Vakil, *A Mathematical Mosaic*, Kelly Publ. Inc., Burlington, Ontario, Canada, 1996.

15. D. Wells, *The Penguin Book of Curious and Interesting Puzzles*, Penguin Books, 1992.

Gregory P. Tollisen   and   Tamás Lengyel
Mathematics Department   Mathematics Department
Occidental College   Occidental College
1600 Campus Road   1600 Campus Road
Los Angeles, CA 90041   Los Angeles, CA
USA   USA
e-mail: `tollisen@oxy.edu`   e-mail: `lengyel@oxy.edu`