A Graphically Motivated, Non-Calculus Derivation of Formulas for Linear Regression

S. M. Zoltek* S. S. Dick†
Department of Mathematical Sciences
and
Honors Program in General Education
George Mason University
Fairfax, Virginia 22030

Abstract

A non-calculus derivation of linear regression formulas is supported by the graphical display of the TI-83 calculator to visualize the minimization of the sum-squared vertical distances.

1 Introduction

The teaching strategies and the examples covered in this paper were developed for a two semester sequence in quantitative problem solving. The courses, *Analysis and Solution of Quantitative Problems I & II*, are part of George Mason University's Plan for Alternative General Education (PAGE) and its new Honors Program in General Education.

The PAGE program (1983–1998) was designed to provide students with a 45-credit, integrative set of courses through which they could complete their general education requirements. Among the goals of the program were developing and using teaching/learning strategies that encouraged student self-learning and discovery.

The Honors Program (established 1997) while adhering to many of the educational goals of the PAGE program, is a highly selective program. The average high school GPA of the entering student is 3.6 and the average SAT score is over 1100. Many of the students have taken AP courses in science and mathematics and about a third express interest in majoring in the sciences. In addition, those not pursuing a science major are better prepared than the "average" student to study the sciences or mathematics.

*Email: szoltek@gmu.edu
†Email: sdick@gmu.edu

In our paper we address a challenge faced by mathematicians teaching a general education mathematics course: *"Excite non-math majors about a quantitative problem and do it at a level that encourages them to explore mathematics."*

This challenge becomes especially relevant when the students body has the potential to major in the sciences, but for now has elected not to do so. For these students, the general education mathematics course must not be a "last look" at mathematics, but rather it should present a world of possibilities and excitement.

Through the development and solution to a specific problem, "guided discovery" is used as a tool to provides students with a chance to "do" mathematics.

While the problem examined is rich enough to provide challenges for gifted students, it leaves room for successes for the less gifted. Our "guided discovery" employs both writing and technology as integral tools in the problem solving process.

Part I of this paper presents a solution to the problem that we believe, while carefully motivated and precise, is accessible to the average college or community college student. "Guided discovery" is used to motivate definitions and make methods of solution plausible. Part II carries the investigation to another level. The student is required to participate in the derivation of equations needed to solve the problem.

Motivation

2 Statement of the Problem and Formulation of a Solution

Through a number of earlier exercises our students have become familiar with the problem solving strategies presented in George Polya's book, *How to Solve It; A New Aspect of Mathematical Method, Doubleday, 1957.* Because of this they are "ready" to undertake the following investigation.

The Problem:

Table 1 lists the cargo capacity and cost of production of four ships. How can this data be used to predict the cost of producing a 5,000 ton capacity cargo ship?

2.1 Understand the Problem

Students are first exposed to the "cargo ship problem" in a homework assignment. This exposure is part of their regular "writing" assignments." (Typical student responses to the sample homework problem appear in italics.)

`Homework #1`

`Understand the Problem:`

1. `What is the unknown?`

   (*The amount it costs to build a 5000 ton capacity cargo ship.*)

2. `What are the data?`

   (*Here they should copy the table or state in complete sentences the capacities and corresponding costs.*)

3. `What is the condition?`

   (*"Not obvious," would be an acceptable response.*)

4. `Draw a figure, introduce notation.`

   See Figure 1.

2.2 Devise a Plan

Many students will note from their graph (see Figure 1) that the data points almost fall on a straight line. Most will believe that if you can find the straight line that comes "closest" to all 4 points, you can extend that line to predict what the cost will be when the value of the independent variable equals 5,000. (Depending on the class, it might be appropriate to review graphing of linear equations and how they relate a given x value to a y value.)

Before they can solve the "cargo ship problem" students will need a better understanding of the phrase, "comes closest to all 4 points."

2.2.1 Simpler Related Problems

Question: Given 3 points, what straight line comes closest to these points? And exactly what do we mean by closest?

Students will find it reasonable to define the closeness of a line to a set of data points to be the sum, S, of the distances from each point in the set to the "candidate" line.

Questions to raise:

- Should we sum the vertical or the perpendicular distances from the points to the "candidate" line? (See Figures 2 and 3.)
- Is simply summing the distances sufficient?

It turns out that we can work with either vertical or perpendicular distances. Statisticians and scientists work with both—one being preferred over the other, depending on the nature of the data being analyzed.

However, computationally, it is much easier to measure the vertical distances. In addition, it is easier to find the line that comes "closest" to a set of data when we speak of "close" in terms of vertical distances.

In Class Exercises:

Exercise 1:

Consider the set of points, $A = \{(-1,-1), (0,1), (1,0)\}$.

Show how an initial guess for a "best-fit" line to A can be made by selecting two points from the data set and connecting the points with a straight line.
Solution: Using the data pair $(-1,-1), (0,1)$, we get $y = 2x + 1$ and using the data pair $(0,1), (1,0)$ we get $y = -x + 1$.

Exercise 2:

a. For each point in the set A, calculate the directed vertical distance to the line $y = 2x + 1$ and sum these directed vertical distances.

b. For each point in the set A, calculate the directed vertical distance to the line $y = -x + 1$ and sum these directed vertical distances.

Exercise 3:

a. Ask students to make their own guess for a "best-fit" line to the data in set A by guessing a slope and an intercept for such a line.

b. For each point in the set A, have students find the directed vertical distance to their line and then have them compute the sum of the distances obtained.
c. Ask for volunteers from the class to state the slope and intercept of their guess and the "sum" associated to it. (You can use the TI-83 program developed below to quickly check their work.)

Exercise 4:

If the class has a strong algebraic preparation, consider computing the sum of the perpendicular distances from the points in the set A to the line $y = 0.5x$. (This line is the least squares line for the data set A.)

Completion of the above exercises points out to students that the process of approximating a best-fit line is time consuming. The visualization provided by a graphing calculator provides "evidence" for rating fits.

The following example points out a problem with defining S to be the sum of the directed vertical distances from the set of data points to the "candidate" line.

Example 1: For each point in the data set $B = \{(-1,-1), (0,0), (1,1)\}$ compute the directed vertical distance to the line $y = x$ and to the line $y = -x$. What is the sum of the directed distances to the line $y = x$? to the line $y = -x$?

Solution: (See Figure 4.) The directed vertical distances from $(-1,-1)$, $(0,0)$, and $(1,1)$ to the line $y = x$ are 0, 0 and 0, respectively. The directed vertical distances from the points $(-1,-1)$, $(0,0)$, and $(1,1)$ to the line $y=-x$ are 2, 0 and -2, respectively. For both "candidate" lines the sum of the directed vertical distances equals 0.

In Class Observations:

- The line $y=-x$ only goes through one of the points, $(0,0)$.
- The line $y=x$ goes through all three data points.
- Our method of measuring "closeness" assigns the same value, $S=0$, to both lines.
- However we define the best-fit line, the line $y=x$ must satisfy this definition for the data set B, since this line "perfectly" fits the data set.

Further Class Discussion:

A possible solution to our dilemma is to take the absolute value of each directed distance *before* adding them together. (In this case we would assign a value of 0 to the line $y=x$ and a value of 4 to the line $y=-x$).

Clearly, adding the absolute values of the directed distances rules out the line $y=-x$ as a "good" fit of our data. Unfortunately, introducing absolute values will complicate the derivation of a formula for finding the best fitting line. We will get around this difficulty by squaring the directed distances before adding them. With this approach the line $y=-x$ will be assigned the value eight $(2^2+0^2+(-2)^2=8)$ and the line $y=x$, again will be assigned the value zero $(0^2+0^2+0^2=0)$.

It can be pointed out to stronger students that taking the absolute value amounts to taking the square root of a square, while simply squaring values will create second degree expressions to which we can apply what we have learned about minimizing quadratic expressions. (At this point in the investigation, this should just be stated and again referred to when expression (1) in Section 2.4.1 is derived below.)

2.2.2 Rating How Well a Straight Line Fits a Set of Data

The above discussion motivates the following method for rating how well a line "fits" a set of data.

Rating the fit of a line:

> Given a line, l, compute the sum the squares of the directed vertical distances between the data points and the line. The value obtained, called $S(l)$, is a measure (rating) of how well the line l fits the data.

The least squares line:

> The line with the smallest value of S is said to have the best fit and is called the *least squares line*. (The method for finding the least squares line is called *linear regression*.)

Example 2: How well does the line, $y = 2x - 1$, fit the set of points, $\{(0,0),(1,1),(2,5)\}$.

Solution: (With as much detail as would be provided in class.)

A. First draw a graph that includes the line and the set of points. (See Figure 5.)

B. Next find the sum of the squares of the directed distances from the data points to the straight line. (See Figure 6.)

Note: If two points, (x_1, y_1) and (x_2, y_2) lie on a vertical line, then $x_1 = x_2$ and the directed distance from (x_1, y_1) to (x_2, y_2) is $y_2 - y_1$.

The vertical line from the point $(0,0)$ to the line $y = 2x - 1$ must cross the line at a point whose x-coordinate is 0. We compute the y-coordinate to be $y = 2 \times 0 - 1 = -1$. So the point of crossing is $(0, -1)$. The directed distance from $(0,0)$ to $(0,-1)$ is $-1 - 0 = -1$. The directed distances from the points $(1,1)$ and $(2,5)$ to the line $y = 2x - 1$ are 0 and -2, respectively. Summing the squares of the directed distances we get $S = (-1)^2 + 0^2 + (-2)^2 = 5$.

2.2.3 The Plan

Our plan, at this point, is to approximate the best-fit line to a set of data by trial and error. First we make an initial guess at a best-fit line, l_1, by guessing a slope and intercept. The line's rating $S(l_1)$ is then computed according to the rule stated in Section 2.2.2. We continue guessing best-fit lines and rating them. From the resulting sequence of lines, $l_1, l_2, \ldots, l_n$, we select the one with the lowest rating. This line will be our "approximation" to the least squares line. (Of course, students should be encouraged to attempt to improve on their guesses.)

After students apply the approximation method to a few examples, we present formulas that give the precise values for the slope and intercept of the least squares line. (See formulas (12).) The students are then asked to compare their guesses with the answers given by the formulas.

2.3 Carry Out the Plan

Students employ the T-83 programmable, graphing calculator as an aid in visualizing and rating the fits of several lines to a given set of data.

2.3.1 A TI-83 Calculator Program for Visualizing and Rating How Well a Line Fits a Set of Data

The process of graphing and rating how well each of several lines fits a set of data is quite tedious. We encourage students to use a program for the TI-83 calculator that simplifies the process.

Our TI-83 program is named AASUMSQR. (See Table 2.) In class, students transfer the program via cable to their calculators. This process moves quite rapidly—a class of 20 students can receive the program in five to ten minutes. (All students taking the Honors/PAGE math sequence are required to own a TI-83 calculator.)

Features of the Program AASUMSQR:

- Data points (x_i, y_i) are accessed from lists L1 and L2, where x_i is stored in the ith position of list L1, and y_i is stored in the ith position of list L2.
- The program sets an appropriate window, plots the data points as a scatter plot, turns on the trace function, and places the trace cursor over the "left-most" data point. (When the trace function is on, the coordinates of a plotted data point are displayed when the trace cursor is placed over the point.)
- The program prompts for the slope and intercept of a candidate least squares line, plots the line on the same screen as the plotted data points, draws vertical lines from the data points to the candidate line and calculates, S, the sum of the squared vertical distances from the points to the candidate line.
- On a single screen, the program displays the guessed values for M and B, the calculated value, S, and the value of S for the previous guess. This makes it easy for students to compare their guesses and "zero in" on a best-fit line.

Application of AASUMSQR to the Cargo Ship Problem

Students use the program, AASUMSQR, to approximate the least squares line for the given cargo ship data and then use the approximated line to estimate the cost of producing a cargo ship that can transport 5,000 tons of cargo.

We find that the hands-on calculator work reinforces the definitions and concepts involved in finding the least squares line. Also, the TI-83's visualization of the fit of a line, makes the estimating process more "real" to students.

For more advanced students, the program sets the stage for viewing S as a function of the two independent variables, slope and intercept.

The Data:

Cargo Capacity (tons)	Average Building Cost ($M)
250	10
500	14
1,000	17
2,000	30

Data entry:

First the student clears lists L1 and L2 and then enters the data points, $\{(x_i, y_i)\}_{i=1}^{n}$ by placing x_i in the ith position of list L1, and y_i in the ith position of list L2.

L1(x)	L2(y)
250	10
500	14
1,000	17
2,000	30

Program execution:

1. The program displays a scatter plot of the data points, $\{(x_i, y_i)\}_{i=1}^{n}$ and turns on the "trace" function. (When the trace function is on, the coordinates of a plotted data point are displayed when the trace cursor is placed over the point.)

2. Pressing ENTER displays, "M=", prompting for the slope of the candidate line. (We assume the student responds with 0.01)

3. Pressing ENTER displays, "B=", prompting for the intercept of the candidate line. (We assume the student responds with 5.)

   Note: A student can make an initial guess at the slope of the least squares line by selecting two of the data points and computing the slope of the line connecting the points. Similar methods can be used to guess at the y intercept of the least squares line.

4. Pressing ENTER displays the line, $y = 0.01x + 5$, the original data points, and vertical lines connecting the data points to the line. (The trace function is again active.)

5. Pressing ENTER we get: (We assume "Float" is set to 4.)

Display 1

M=	
	0.0100
B=	
	5.0000
OLD S	NEW S
	-1.0000
	51.2500

Display 1 shows the slope and intercept chosen, the value of S for the previously chosen line (Old S), and the value of S for the line just chosen (New S=51.25). Since this is the first candidate line chosen, OLD S is set to, -1, to indicate there was no previous guess made.

6. Pressing ENTER redisplays the scatter plot with the superimposed candidate line. (The student will notice that the line displayed is below all of the data points. The value for NEW S is 51.25.)

7. Pressing ENTER, again prompts with, "M=", for the slope of the next candidate line. (We assume the student again enters 0.01 since the slope seems about right.)

8. Pressing ENTER, again prompts with, "B=". (The old value for B, which is still displayed on the screen, was 5. We assume the student responds with 8.)

9. Pressing ENTER displays the new line, $y = 0.01x + 8$, superimposed on the scatter plot of data points. (The old line is not displayed.)

10. Pressing ENTER brings up Display 2, below, with the same format as Display 1, but showing the slope and intercept of the new candidate line, and the value of S for the previous and current candidate lines.

Display 2

M=	
	0.0100
B=	
	8.0000
OLD S	NEW S
	51.2500
	6.2500

Notice that the new value of S, 6.25, is much smaller than the previous value, 51.25, which shows the student that he or she is has made a much better estimate.

At this point students are encouraged to use the best line they have found thus far to make a guess at the solution to the cargo ship problem. If the "best" line obtained by "guessing," aided by the TI-83 program, is $y = 0.01x + 8$, then the best guess at the cost of the ship will be $y = (0.01)(5000) + 8 = 58$, or $58 million.

Next, we would provide the equations for the slope and intercept of the least squares line. (See equations (12).)

Then students would be asked to compare their approximation to the cost of producing a 5,000 ton capacity cargo ship with the cost predicted using the least squares line. (See Section 2.4.3.)

Derivation

2.4 Carry Out the Plan—Extended Version

The visualization and computation provided by the program AASUMSQR has set the stage for the derivation of formulas (12) for the least squares line.

Students should be reminded that using the program AASUMSQR to rate a line involves inputting the line's slope, m, and intercept, b. So the output, S, depends on inputs m and b. Specifically, students should be directed to Display 1 and Display 2, above, and asked to state the program's output when $(0.01, 5)$ and $(0.01, 8)$, respectively, are the inputs. (Though we assume our Honors Program students are familiar with the concept of a function, we assign a detailed homework assignment that reviews functions of a single variable and introduces functions of two variables.)

Use of the program, together with the above discussion should make it clear that finding the least squares line involves minimizing a function of two variables. The next step is to write down the exact expression that needs to be minimized.

2.4.1 The Algebraic Expression to be Minimized

A series of homework exercise leads to the general expression (value) for S when the set of data points and line are given by $\{(x_1, y_1), (x_2, y_2), (x_3, y_3)\}$ and $y = mx + b$, respectively. Figure 7 is included with the assignment.

The solution is:

$$S = (y_1 - mx_1 - b)^2 + (y_2 - mx_2 - b)^2 + (y_3 - mx_3 - b)^2 \qquad (1)$$

Depending on the mathematical preparation of the class, one might decide to restrict the "general" discussion to that of finding the least squares line for data sets consisting of three ponts. (When there are only two points, the answer is to just draw the line connecting the two points.)

Whether or not the class is algebraically up to the most general approach, we always do the computations for three points and then assign as homework the task of repeating the computations for four points. After the homework is reviewed, we restate each computation in the setting of n points. This is preceded by a careful discussion of sigma notation.

Keeping the above instructional notes in mind, we will continue deriving the formulas for the least squares line with the assumption that our data set has n points.

Note: When the data set has n points the expression for S becomes:

$$\begin{aligned} S &= (y_1 - mx_1 - b)^2 + (y_2 - mx_2 - b)^2 + (y_3 - mx_3 - b)^2 \\ &\quad + \cdots + (y_n - mx_n - b)^2 \\ &= \sum_{i=1}^{n} (y_i - mx_i - b)^2 \end{aligned} \tag{2}$$

2.4.2 A Close Examination of the Expression for S

In class, the expression for S is examined in more detail.

1. The expression for S in (2) is expanded and it is observed that the highest degree of any variable is 2.
2. Students are asked, "What do you know about quadratic expressions in one variable that might be useful in finding the minimum value of S?"
3. If necessary, a review follows that produces or states the following two facts:
   (a) The parabola,
   $$y = ux^2 + vx + w \tag{3}$$
   is concave up if and only if $u > 0$.
   (b) The minimum y value for parabola, (3), occurs when $x = -v/(2u)$.

Building on students' understanding of parabolas, we rewrite the expression for S as a quadratic in powers of b. (For this computation we consider m to be a fixed, constant value.)

After expanding the expression for S and collecting the coefficients of b^2 and b, we get

$$\begin{aligned} S &= nb^2 + \left[2m\sum_{i=1}^{n} x_i - 2\sum_{i=1}^{n} y_i\right] b \\ &\quad + \left[\sum_{i=1}^{n} y_i^2 - 2m\sum_{i=1}^{n} x_i y_i + m^2 \sum_{i=1}^{n} x_i^2\right] \end{aligned} \tag{4}$$

From equation (4) we see the coefficient of b^2 is n and the coefficient of b is $[2m\sum_{i=1}^n x_i - 2\sum_{i=1}^n y_i]$.

Comparing (3) and (4) we have:

$$u = n \text{ and } v = \left[2m\sum_{i=1}^n x_i - 2\sum_{i=1}^n y_i\right]$$

From the review on properties of parabolas and noting that n, the coefficient of b^2, is positive, it follows that for each fixed m, S will be smallest when

$$\begin{aligned} b &= -v/(2u) = -\frac{2m\sum_{i=1}^n x_i - 2\sum_{i=1}^n y_i}{2n} \\ &= \frac{\sum_{i=1}^n y_i - m\sum_{i=1}^n x_i}{n} \\ &= \frac{\sum_{i=1}^n y_i}{n} - \frac{m\sum_{i=1}^n x_i}{n} \end{aligned} \tag{5}$$

Note: If we set $\overline{y} = \left(\sum_{i=1}^n y_i\right)/n$ and $\overline{x} = \left(\sum_{i=1}^n x_i\right)/n$, then equation (5) becomes

$$b = \overline{y} - m\overline{x} \qquad \text{or} \qquad \overline{y} = m\overline{x} + b \tag{6}$$

Equation (6) tells us that the least squares line must go through the point $(\overline{x}, \overline{y})$

So far we have found a formula, (6), for the y-intercept of the least squares line of a set of data. However, our formula requires that we know the slope, m, of the line.

The formula, $\overline{y} = m\overline{x} + b$, is a linear equation relating m and b. Next we will find another linear equation in m and b. The simultaneous solution of these two linear equations in the two unknowns (m and b) will determine uniquely the least squares line.

The technique for finding a second linear equation involving m and b is similar to that for finding equation (5).

We again expand the expression for S, but this time we collect coefficients of m^2 and m.

We get:

$$\begin{aligned} S &= \left[\sum_{i=1}^n {x_i}^2\right] m^2 + \left[2b\sum_{i=1}^n x_i - 2\sum_{i=1}^n x_1 y_i\right] m \\ &\quad + \left[\sum_{i=1}^n {y_i}^2 - 2b\sum_{i=1}^n y_i + nb^2\right] \end{aligned} \tag{7}$$

Again from the review of the properties of parabolas and noting that the coefficient of m^2, $\sum_{i=1}^n {x_i}^2$, is a sum of squares and so is positive, it follows that for each b, S will be smallest when

$$
\begin{aligned}
m &= -v/(2u) = -\frac{2b\sum_{i=1}^{n} x_i - 2\sum_{i=1}^{n} x_1 y_i}{2\sum_{i=1}^{n} {x_i}^2} \\
&= \frac{\sum_{i=1}^{n} x_1 y_i - b\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} {x_i}^2} \qquad (8)
\end{aligned}
$$

Equation (8) can be rewritten as

$$
m\sum_{i=1}^{n} {x_i}^2 = \sum_{i=1}^{n} x_i y_i - b\sum_{i=1}^{n} x_i \qquad (9)
$$

Collecting coefficients of m and b, equations (6) and (9) can be written as

$$
\left\{ \begin{array}{rcrcl}
(\sum_{i=1}^{n} {x_i}^2)\mathrm{m} & + & (\sum_{i=1}^{n} x_i)\, b & = & \sum_{i=1}^{n} x_i y_i \\
\overline{x}\ \mathrm{m} & + & b & = & \overline{y}
\end{array} \right\} \qquad (10)
$$

The equations in (10) form a set of two linear equations in the two unknowns, m and b.

Solving the second equation of (10) for b (see (6)) and substituting in the first equation of (10), we get

$$
m = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} {x_i}^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \qquad (11)
$$

Together equations (6) and (11) give us a method for finding the least squares line.

Formulas for the Slope and Intercept of the Least Squares Line:

$$
\left\{ \begin{array}{rcl}
m & = & \dfrac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} {x_i}^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \\
b & = & \overline{y} - m\overline{x}
\end{array} \right\} \qquad (12)
$$

2.4.3 Solution to the Cargo Ship Problem

Using our data set (see Table 1) and doing the one-variable statistics with a TI-83 calculator we get:

$\sum_{i=1}^{n} x_i y_i = 86,500$ $\quad \sum_{i=1}^{n} x_i = 3,750$ $\quad \left(\sum_{i=1}^{n} x_i\right)^2 = 14,062500$

$\sum_{i=1}^{n} y_i = 71$ $\quad \sum_{i=1}^{n} {x_i}^2 = 5,312,500$

$\overline{x} = 937.5$ $\quad \overline{y} = 17.75$

Substitution into (12) yields $m = 0.0111$ and $b = 7.3478$. (In calculating b, the full accuracy of the calculator was used, not just the value, 0.0111.)

Therefore the least squares line is $y = 0.0111x + 7.3478$ and the best guess for the cost of a cargo ship with a cargo capacity of 5,000 pounds is $y = (0.0111)(5000) + 7.3478 = 62.8478$ million dollars.

3 Summary

We have found that the amount of student preparation needed before students can work through the derivation of equations such as (1), (4) and (11) differs greatly from class to class and student to student. Because of this we assign a comprehensive review exercise set that prepares students with the algebraic skills needed for the type of computations they will be performing. Each set is assigned prior to the day on which the algebraic skill will be needed.

Table 1: (Capacity, Cost)

Cargo Capacity (tons)	Average Building Cost ($M)
250	10
500	14
1,000	17
2,000	30
5,000	?

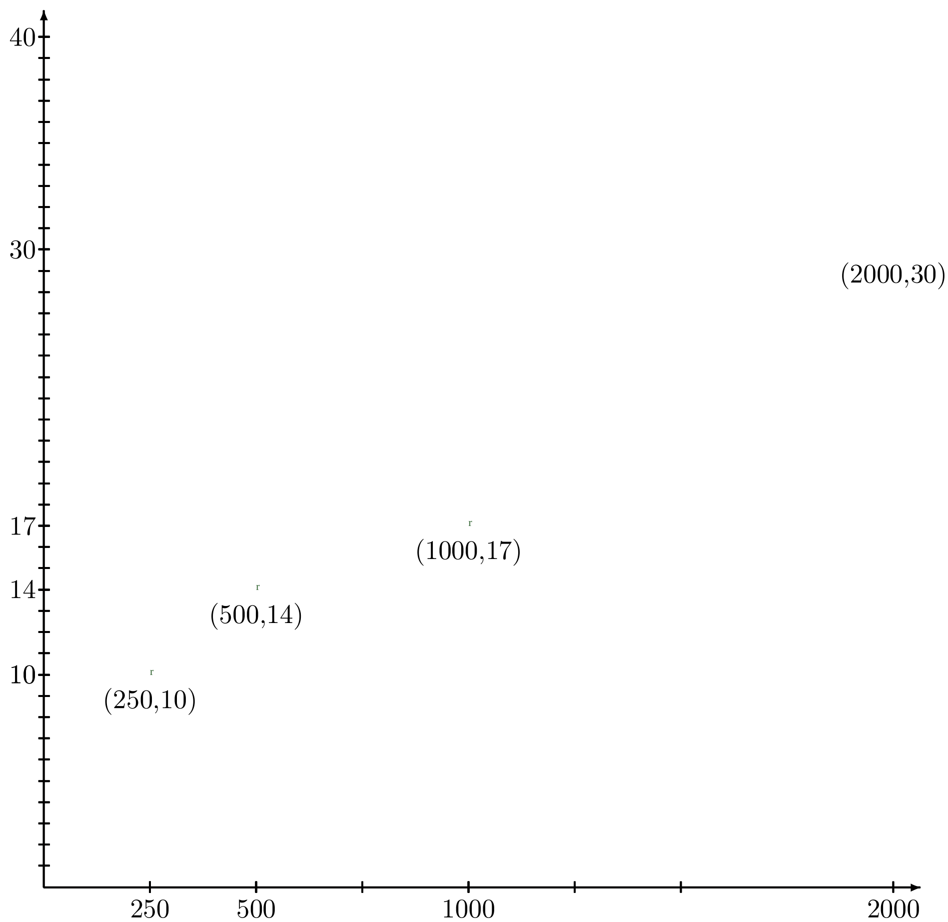

Figure 1:

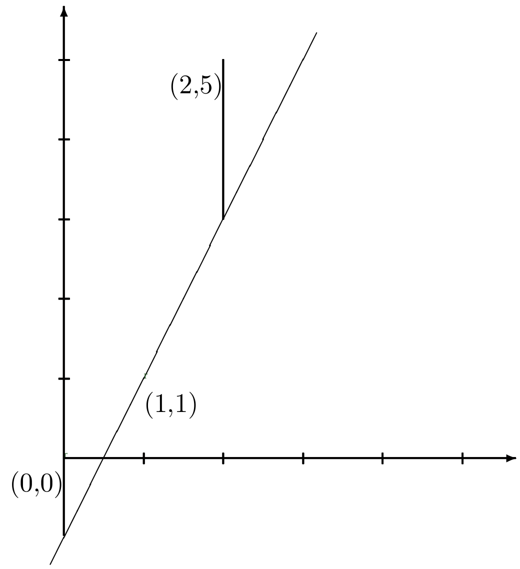

Figure 2: The Vertical Distances

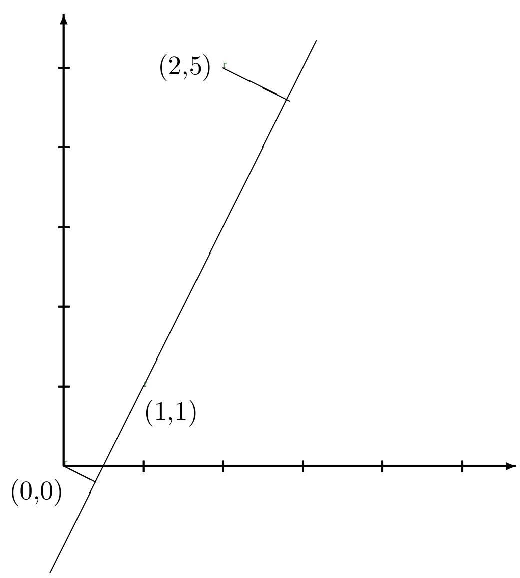

Figure 3: The Perpendicular Distances

Table 2: The TI-83 Program AASUMSQR

```
PROGRAM:AASUMSQR
:FnOff
:PlotsOff
:min(0,min(L1)-.05(max(L1)-(min(L1)))→ Xmin
:max(0,max(L1)+.05(max(L1)-min(L1))→Xmax
:min(0,min(L2)-.05(max(L2)-min(L2)))→Ymin
:max(0,max(L2)+.05(max(L2)-min(L2))→Ymax
:Plot1(Scatter,L1,L2,■)
:DispGraph
:Trace
:-1→E
:Lbl A
:Prompt M,B
:FnOff
:"MX+B"→Y1
:1→J
:While J≤dim(L1)
:Line(L1(J),L2(J),L1(J),Y1(L1(J)))
:1+J→J
:End
:1→I
:0→D
:While I≤dim(L1)
:(Y1(L1(I))-L2(I))^2+D→D
:I+1→I
:End
:Trace
:Disp "M=",M
:Disp "B=",B
:Disp "OLD S, NEW S"
:Disp E, D
:D→E
:Pause
:DispGraph
:Trace
:Goto A
```

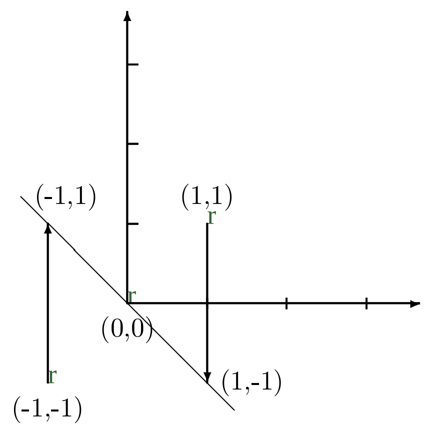

Figure 4:

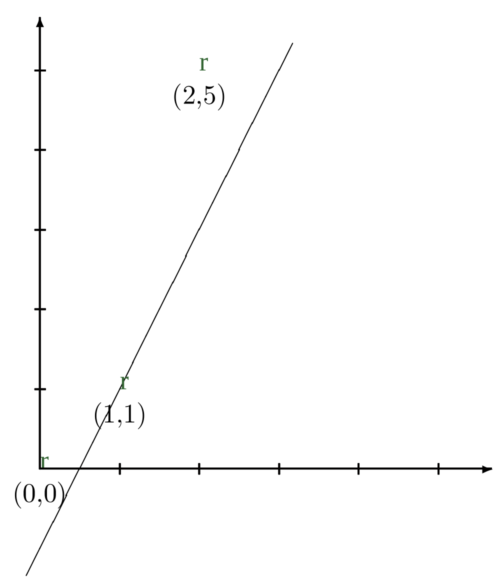

Figure 5:

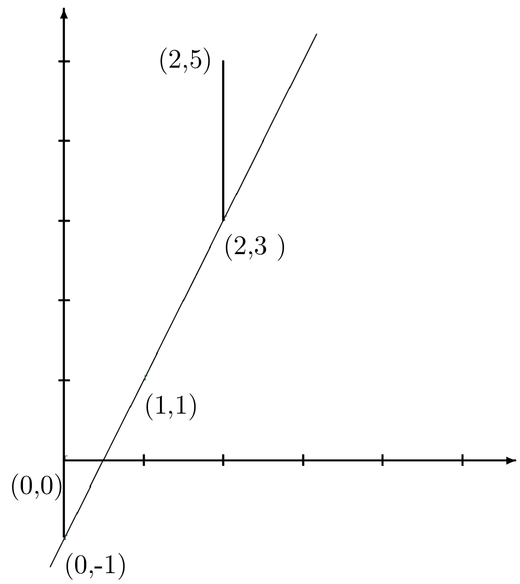

Figure 6: Computing The Vertical Distances

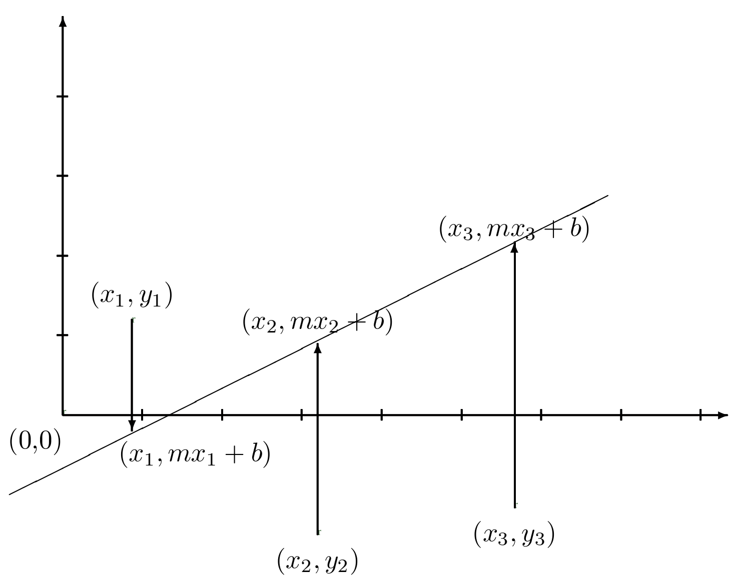

Figure 7: Vertical Distances to a Line From Three Arbitrary Points