Some Geometric Interpretations of The Total Distances Among Curves and Surfaces

Wei-Chi Yang
wyang@radford.edu
Department of Mathematics and Statistics
Radford University
Radford, VA 24142
USA

Abstract

In this paper, we explore some examples in findinging the global values of a total squared distances among non-intersecting curves in the plane and surfaces in the space. We hope the geometric interpretations will help studdents appreciate the use of Lagrange Multipliers Method and concepts learned from Linear Algebra.

1 Introduction

Throughout this paper, curves are non-intersecting in the plane $\mathbb{R}^2$ and surfaces are non-intersecting in the space $\mathbb{R}^3$. In Sections 2 and 3, we give motivations and examples of finding the global squared distance between two curves or surfaces. Finding the global squared distance between two nonconvex surfaces has been researched in [2], but the method is applied there only to some special cases. For general curves and surfaces, the motivations described in Sections 2 and 3 are still crucial for students to comprehend how Lagrange Multipliers can be viewed geometrically. In Sections 4 and 5, we extend this idea further: We give three disjoint curves, C_1, C_2 and C_3 respectively, and we want to find the global value of the total squared distances from C_1 to C_2 and C_1 to C_3. We extend the idea from 2D to 3D on convex surfaces. We are given four **convex** surfaces S_1, S_2, S_3 and S_4, we would like to find the the global value of the total squared distances from S_1 to S_2, S_1 to S_3, and S_1 to S_4. The method of finding the extremum of a total squared distances is an application of Theorem 4 and is proved in Corollary 5. Theorem 4 is a generalization of Corollary 5. For simplicity, we will calculate the squared distance $|\mathbf{x}-\mathbf{y}|^2$ instead of $|\mathbf{x}-\mathbf{y}|$, where $\mathbf{x}$ and $\mathbf{y}$ are in $\mathbb{R}^2$ or $\mathbb{R}^3$.

2 Motivations in 2-D

Let curves C_1 and C_2 be two non-intersecting curves of forms $f(x,y) = 0$ (or $y = f(x)$) and $g(x,y) = 0$ (or $y = g(x)$) respectively. If we restrict the ranges for x and y for both curves to be closed and bounded sets, the global maximum and minimum exist by the Extreme Value Theorem. (We remark that if both C_1 and C_2 are lines, they have to be parallel in our discussion here.) By following the idea we described in the video clip (see [6]), which uses the software package [ClassPad] for demonstration), we note that a necessary conditions for the squared distance AB, where $A \in C_1$ and $B \in C_2$, to be either the minimum or maximum is that:

$$\overrightarrow{AB} \text{ is perpendicular to the tangent line at } A \text{ and}$$
$$\overrightarrow{AB} \text{ is perpendicular to the tangent line at } B. \tag{1}$$

We demonstrate this observation using the following example.

Example 1 *Let* $f(x) = \sin(x) + 5$ *and* $g(x) = -\cos(x-1)$. *We discuss the maximum and minimum squared distances between* $y = f(x)$ *and* $y = g(x)$ *in the closed interval of* $[-2, 5] \times [-4, 8]$.

Case 1. By using (1) and with the help of Maple (see [7]), we found that when

$$\begin{aligned} x_1 &= 1.685610214 \text{ and } f(x_1) = 5.993416123, \\ x_2 &= .8851861123, \text{ and } g(x_2) = -.9934161229. \end{aligned}$$

The squared distance between $(x_1, f(x_1))$ and $(x_2, g(x_2)$ is about 49.45650357. We sketch the graphs of $y = f(x)$, and $y = g(x)$, the line segment connecting $(x_1, f(x_1))$ and $(x_2, g(x_2)$, the tangent line at $x = x_1$ and the tangent line at $x = x_2$ together as follows. This case turns out to produce the global maximum squared distance.

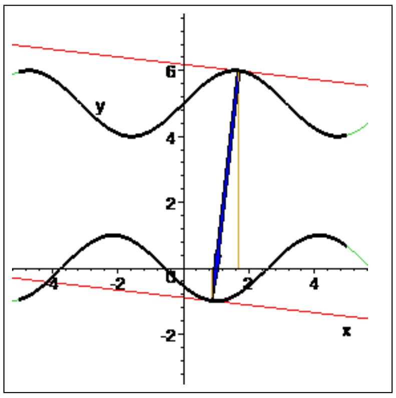

Figure 1.Global Maximum Squared Distance in$[-2, 5] \times [-4, 8]$

Case 2. When

$$\begin{aligned} x_5 &= -1.684808257 \text{ and } f(x_5) = 4.006492323, \\ x_6 &= -2.027580723, \text{ and } g(x_6) = .9935076771, \end{aligned}$$

the squared distance between $(x_5, f(x_5))$ and $(x_6, g(x_6)$ is about 9.195569440. We sketch the graphs of $y = f(x)$, and $y = g(x)$, the line segment connecting $(x_5, f(x_5))$ and $(x_6, g(x_6)$, tangent line at

$x = x_5$ and tangent line at $x = x_6$ together as shown in **Figure 2**.

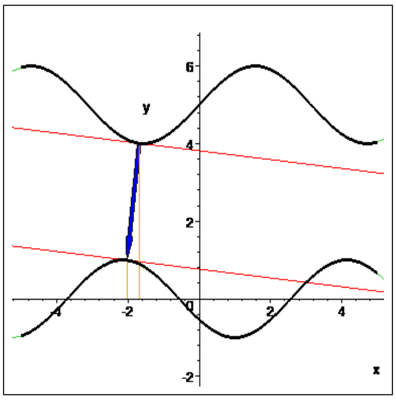

Figure 2. Global Minimum Squared Distance in $[-2, 5] \times [-4, 8]$

This turns out to be the case producing the global minimum squared distance.

Remark: We note that (1) only produces the relative extremum for the squared distance function, we could get a lot more possibilities than two cases we mentioned above. For example, the following set of points satisfies (1):

$$\begin{aligned} x_3 &= 3.421927911 \text{ and } f(x_3) = 4.723322166, \\ x_4 &= -.8511315841, \text{ and } g(x_4) = .2766778338. \end{aligned}$$

The squared distance between $(x_3, f(x_3))$ and $(x_4, g(x_4)$ is roughly 38.03168327 only represents a relative extremum for the squared distance function. We sketch the graphs of $y = f(x)$, and $y = g(x)$, the line segment connecting $(x_3, f(x_3))$ and $(x_4, g(x_4)$, tangent line at $x = x_3$ and tangent line at $x = x_4$ together as follows:

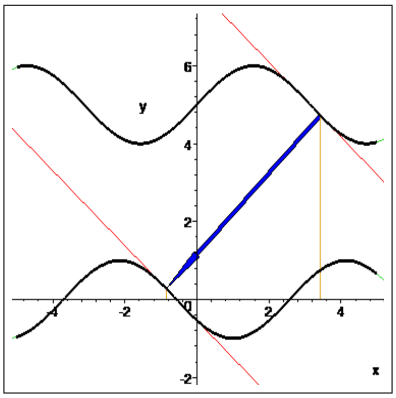

Figure 3. Relative Extremum Squared Distance in $[-2, 5] \times [-4, 8]$

Exercise 1. Verify that the following set of points produces only a relative extremum for Example 1.

$$\left\{\begin{bmatrix} -.4955520442 \\ f(-.4955520442) \end{bmatrix}, \begin{bmatrix} 3.066348371 \\ g(3.066348371) \end{bmatrix}\right\},$$
$$\left\{\begin{bmatrix} 1.030769192 \\ f(1.030769192) \end{bmatrix}, \begin{bmatrix} 3.066348371 \\ g(3.066348371) \end{bmatrix}\right\},$$
$$\left\{\begin{bmatrix} 1.685610214 \\ f(1.685610214) \end{bmatrix}, \begin{bmatrix} .8851861123 \\ g(.8851861123) \end{bmatrix}\right\}.$$

3 Motivations in 3-D

Let f and g be continuously differentiable functions in their respective closed and bounded domains (which are subsets of $\mathbb{R}^3$). Here we describe how we find the relative extremum squared distance between two convex surfaces satisfying

$$f(x,y,z) = 0 \text{ and } g(x,y,z) = 0. \tag{2}$$

Following what we have described in two dimensional case, we see that if $A = (x_1, x_2, x_3)$ is on the surface $f(x,y,z) = 0$ and $B = (y_1, y_2, y_3)$ is on the surface $g(x,y,z) = 0$, a necessary condition for finding such extremum distance is to have

$$\begin{gathered} \overrightarrow{AB} \text{ is parallel to the normal vector of the tangent plane at } A \text{ and} \\ \overrightarrow{AB} \text{ is parallel to the normal vector of the tangent plane at } B. \end{gathered} \tag{3}$$

More specifically, if we use $\nabla f = (f_x, f_y, f_z)$ and $\nabla g = (g_x, g_y, g_z)$ to denote the gradients of f and g respectively, the following conditions should be met.

$$\begin{gathered} \overrightarrow{AB} = \lambda_1 (\nabla f) \text{ at } A, \\ \overrightarrow{AB} = \lambda_2 (\nabla g) \text{ at } B, \\ f(x_1, x_2, x_3) = 0, \text{ and} \\ g(y_1, y_2, y_3) = 0. \end{gathered} \tag{4}$$

Simply put, the vectors (∇f) at A and (∇g) at B are parallel. The above set of equations in (4) is equivalent to the followings:

$$y_1 - x_1 = \lambda_1 (f_x)_{(x_1,x_2,x_3)}, \tag{5}$$
$$y_2 - x_2 = \lambda_1 (f_y)_{(x_1,x_2,x_3)}, \tag{6}$$
$$y_3 - x_3 = \lambda_1 (f_z)_{(x_1,x_2,x_3)}, \tag{7}$$
$$y_1 - x_1 = \lambda_2 (g_x)_{(y_1,y_2,y_3)}, \tag{8}$$
$$y_2 - x_2 = \lambda_2 (g_y)_{(y_1,y_2,y_3)}, \tag{9}$$
$$y_3 - x_3 = \lambda_2 (g_z)_{(y_1,y_2,y_3)}, \tag{10}$$
$$f(x_1, x_2, x_3) = 0 \text{ and } g(y_1, y_2, y_3) = 0. \tag{11}$$

It is easy to verify that equations (5)-(11) produce the identical result if we apply the Lagrange Multiplier Method in solving this problem. In other words, we want to minimize or maximize the squared distance $|\mathbf{x}-\mathbf{y}|^2$, which is subject to both $f(\mathbf{x})=0$ and $g(\mathbf{y})=0$. Specifically, if we write

$$L(\mathbf{x},\mathbf{y},\lambda_1,\lambda_2)=|\mathbf{x}-\mathbf{y}|^2+\lambda_1 f(\mathbf{x})+\lambda_2 g(\mathbf{y}) \tag{12}$$

or

$$\begin{aligned} L(x_1,x_2,x_3,y_1,y_2,y_3,\lambda_1,\lambda_2) &= (x_1-y_1)^2+(x_2-y_2)^2+(x_3-y_3)^2+ \\ &\quad \lambda_1 f(x_1,x_2,x_3)+\lambda_2 g(y_1,y_2,y_3); \end{aligned} \tag{13}$$

Then it follows from the Lagrange Multiplier Method that a necessary condition to achieve the critical distance is to have

$$\nabla L = 0. \tag{14}$$

We demonstrate this effect by using the following example.

Example 2 *Let $f(x,y,z)=\sin x\cos y-2-z$ and $g(x,y,z)=x^2+y^2-z$. If we restrict the domain to be $[-2,2]\times[-2,2]\times[-4,4]$ for both functions, we are interested in finding the global minimum of the squared distance between $f(x,y,z)=0$ and $g(x,y,z)=0$.*

By using equation (12) and setting $\nabla L=0$, the computation from Maple shows one of the following solutions below:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} .9776334541 \\ 0 \\ -1.170823173 \end{bmatrix}, \text{ and}$$
$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} .2794931976 \\ 0 \\ .07811644752 \end{bmatrix}.$$

The squared distance $|\mathbf{x}-\mathbf{y}|^2$ is 2.047249995. It follows from our computation that

$$(\nabla f) \text{ at } (x_1,x_2,x_3) = (.5589863953, 0, -1), \text{ and} \tag{15}$$
$$(\nabla g) \text{ at } (y_1,y_2,y_3) = (.5589863952, 0, -1]), \tag{16}$$

which shows that (∇f) at (x_1,x_2,x_3) is (almost) identical to (∇g) at (y_1,y_2,y_3) (subject to some computation round off error). The graph below shows the solution set will achieve the global minimum distance for these two surfaces.

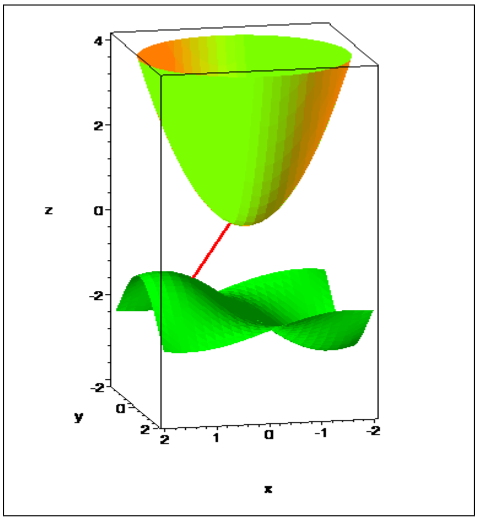

Figure 4. The shortest distance between two surfaces in a closed and bounded domain

We extend the ideas described in Examples 1 and 2 to the following scenario. First, we denote the distance between two points A and B in the space by $\|AB\|$. We are given four convex surfaces in the space, represented by the **orange, yellow, blue and purple** surfaces which we call them S_1, S_2 S_3 and the S_4 respectively. We want to find points A, B, C and D on S_1, S_2, S_3 and S_4 respectively so that the $\|AB\| + \|AC\| + \|AD\|$ achieves its minimum.

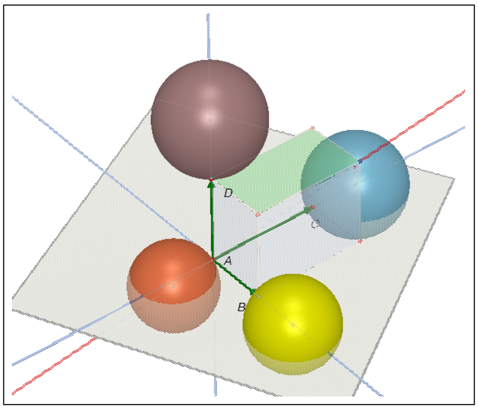

Figure 5. The shortest total squared distances among four convex surfaces

It follows from our discussions in 2-D and 3-D above that the followings should be met.

1. The vector AB should be parallel to the normal vector of the surface S_2 at B. This is equivalent to
$$\overrightarrow{AB} = \lambda_2 \left(\nabla S_2 \text{ at } B\right) \text{ for some } \lambda_2. \tag{17}$$

2. The vector AC should be parallel to the normal vector of the surface S_3 at C. This is equivalent to
$$\overrightarrow{AC} = \lambda_3 \left(\nabla S_3 \text{ at } C\right) \text{ for some } \lambda_3. \tag{18}$$

3. The vector AD should be parallel to the normal vector of the surface S_4 at D.This is equivalent to
$$\overrightarrow{AD} = \lambda_4 \left(\nabla S_2 \text{ at } D\right) \text{ for some } \lambda_4. \tag{19}$$

4. To achieve the minimum of $\|AB\| + \|AC\| + \|AD\|$ we should also place the points A so that the normal vector of S_1 at A is a linear combination of $\overrightarrow{AB}$, $\overrightarrow{AC}$ and $\overrightarrow{AD}$. This is equivalent to say we can find λ_1 so that

$$\lambda_1 \left(\nabla S_1 \text{ at } A\right) = \lambda \left({}_2\nabla S_2 \text{ at } B\right) + \lambda_3 \left(\nabla S_3 \text{ at } C\right) + \lambda_4 \left(\nabla S_2 \text{ at } D\right). \tag{20}$$

We shall proceed proving our observations in Corollary 5 of Section 4.

4 Applications of Lagrange Theorem

We remind readers about the Lagrange Multiplier Method in Theorem 3 (without proof), which can be found in many textbooks. We shall see that finding the global values of a total squared distances is a special case of the Theorem 4.

Theorem 3 *We assume that f, g are continuously differentiable: $R^n \to R$. Suppose that we want to maximize or minimize a function of n variables $f(\mathbf{x}) = f(x_1, x_2, ..., x_n)$ for $\mathbf{x} = (x_1, x_2, ..., x_n)$ subject to p constraints $g_1(\mathbf{x}) = c_1, g_2(\mathbf{x}) = c_2, ...,$ and $g_p(\mathbf{x}) = c_p$. The necessary condition of finding the relative maximum or minimum of $f(\mathbf{x})$ subject to the constraints $g_1(\mathbf{x}) = c_1, g_2(\mathbf{x}) = c_2, ...,$ and $g_p(\mathbf{x}) = c_p$ that is not on the boundary of the region where $f(\mathbf{x})$ and $g_i(\mathbf{x})$ are defined can be found by solving the system*

$$\frac{\partial}{\partial x_i}\left(f(\mathbf{x}) + \sum_{j=1}^{p} \lambda_j g_j(\mathbf{x})\right) = 0, \; 1 \le i \le n, \tag{21}$$
$$g_j(\mathbf{x}) = c_j, \; 1 \le j \le p. \tag{22}$$

We write $\nabla f(\mathbf{x}) = \left(\frac{\partial}{\partial x_1} f(\mathbf{x}), \frac{\partial}{\partial x_2} f(\mathbf{x}), ..., \frac{\partial}{\partial x_n} f(\mathbf{x})\right)$. If $x = \mathbf{x}_0$ is an extremum for above system, then
$$\nabla f(\mathbf{x}_0) = \sum_{j=1}^{p} \lambda_j \nabla g_j(\mathbf{x}_0). \tag{23}$$

We now consider a slight variation here. We remark that the Theorem 4 below is a generalization of Corollary 5. However, Corollary 5 is inspired by finding the global value of a squared distance function.

We assume that $f : R^{np} \to R, g_i : R^n \to R, i = 1, 2...p$, are continuously differentiable in their respective domains. Our objective is to maximize or minimize the function

$$\begin{aligned} f(\mathbf{x}_1, \mathbf{x}_2, , , , \mathbf{x}_p) &= f(x_1^1, x_2^1, ..., x_n^1, x_1^2, x_2^2, ..., x_n^2, ...x_1^p, x_2^p, ..., x_n^p) \\ \text{for } \mathbf{x}_i &= (x_1^i, x_2^i, ..., x_n^i),\ i = 1, 2, ...p \end{aligned} \tag{24}$$

subject to p constraints $g_1(\mathbf{x}_1) = c_1, g_2(\mathbf{x}_2) = c_2, ...$, and $g_p(\mathbf{x}_p) = c_p$.

Theorem 4 *A necessary condition of finding the relative maximum or minimum of $f(\mathbf{x}_1, \mathbf{x}_2, , , , \mathbf{x}_p)$ subject to the constraints $g_1(\mathbf{x}_1) = c_1, g_2(\mathbf{x}_2) = c_2, ...$, and $g_p(\mathbf{x}_p) = c_p$ that is not on the boundary of the region where $f(\mathbf{x}_1, \mathbf{x}_2, , , , \mathbf{x}_p)$ and $g_i(\mathbf{x}_i)$ are defined can be found by solving the system*

$$\frac{\partial}{\partial \mathbf{x}_i}\left(f(\mathbf{x}_1, \mathbf{x}_2, , , , \mathbf{x}_p) + \sum_{j=1}^{p} \lambda_j g_j(\mathbf{x}_j)\right) = 0,\ 1 \le i \le p, \tag{25}$$

$$g_j(\mathbf{x}_j) = c_j,\ 1 \le j \le p. \tag{26}$$

If

$$\sum_{i=1}^{p} \frac{\partial}{\partial \mathbf{x}_i}\left(f(\mathbf{x}_1, \mathbf{x}_2, , , , \mathbf{x}_p)\right) = 0, \tag{27}$$

and $\mathbf{x} = \mathbf{x}_0 = (\mathbf{x}_1^*, \mathbf{x}_2^*, , , , \mathbf{x}_p^*)$ is an extremum, then we have

$$\sum_{j=1}^{p} \lambda_j \frac{\partial}{\partial \mathbf{x}_i}\left(g_j(\mathbf{x}_j^*)\right) = 0. \tag{28}$$

Proof: We observe that $\frac{\partial}{\partial \mathbf{x}_i} f(\mathbf{x}_1, \mathbf{x}_2, , , , \mathbf{x}_p) = -\lambda_i \frac{\partial}{\partial \mathbf{x}_i} g_i(\mathbf{x}_i), i = 1, 2, ...p$.

If $\sum_{i=1}^{p} \frac{\partial}{\partial \mathbf{x}_i}\left(f(\mathbf{x}_1, \mathbf{x}_2, , , , \mathbf{x}_p)\right) = 0$, and $\mathbf{x} = \mathbf{x}_0 = (\mathbf{x}_1^*, \mathbf{x}_2^*, , , , \mathbf{x}_p^*)$ is an extremum, then

$$\sum_{j=1}^{p} \lambda_j \frac{\partial}{\partial \mathbf{x}_i}\left(g_j(\mathbf{x}_j^*)\right) = 0. \tag{29}$$

Corollary 5 *If the total squared distances function $f(\mathbf{x}_1, \mathbf{x}_2, , , , \mathbf{x}_p) = |\mathbf{x}_1 - \mathbf{x}_2|^2 + |\mathbf{x}_1 - \mathbf{x}_3|^2 + ... |\mathbf{x}_1 - \mathbf{x}_p|^2$* has a global value, where $\mathbf{x}_i = (x_1^i, x_2^i, ..., x_n^i), i = 1, 2, ...p$, *subject to p constraints*

$$g_1(\mathbf{x}_1) = c_1, g_2(\mathbf{x}_2) = c_2, ..., \text{ and } g_p(\mathbf{x}_p) = c_p, \tag{30}$$

at $\mathbf{x}_0 = (x_1^*, x_2^*, , , , x_p^*)$ in its closed and bounded domain. Then we can find coefficients, $\lambda_j, j = 1, 2, ...p$, such that

$$\lambda_1 \nabla g_1(\mathbf{x}_0) = \sum_{j=2}^{p} \lambda_j \nabla g_j(\mathbf{x}_0). \tag{31}$$

Proof: We set

$$L(\mathbf{x}_1, \mathbf{x}_2, , , , \mathbf{x}_p) = |\mathbf{x}_1 - \mathbf{x}_2|^2 + |\mathbf{x}_1 - \mathbf{x}_3|^2 + ... |\mathbf{x}_1 - \mathbf{x}_p|^2 + \sum_{j=1}^{p} \lambda_j g_j(\mathbf{x}_j). \tag{32}$$

By setting $\nabla L = 0$, we observe that $\{\mathbf{x}_1, \mathbf{x}_2, , , , \mathbf{x}_p\}$ are such that $g_1(x_1) = c_1, g_2(x_2) = c_2$, ..., and $g_p(x_p) = c_p$. Furthermore,

$$\begin{aligned} 2p\mathbf{x}_1 - 2\sum_{j=2}^{p} \mathbf{x}_j &= -\lambda_1 \nabla g_1 && (33) \\ -2(\mathbf{x}_1 - \mathbf{x}_2) &= \lambda_2 \nabla g_2 && (34) \\ & \dots && (35) \\ -2(\mathbf{x}_1 - \mathbf{x}_p) &= \lambda_p \nabla g_p, \text{ and} && (36) \end{aligned}$$

The proof follows directly by adding up the left hand sides and right hand sides of equations (33)-(36) separately.

Remark: We note that if we are given $n+1$ vectors in $\mathbb{R}^n$, they form a linearly dependent set. The Corollary 5 says that finding an extremum for the the total distance squared function in $p = n+1$ case, the vector $\nabla g_1(\mathbf{x}_0)$ can be written as a linear combination of the remaining gradient vectors at the critical point.

We give some examples to demonstrate the use of Theorem 4 and Corollary 5 in the next section.

5 Examples in 2-D and 3-D

First we describe how Corollary 5 can be used in 2-D and 3-D cases.

Example 5 *Let $g_1(x_1, y_1) = \sin x_1 - y_1$, $g_2(x_2, y_2) = x_2^2 - y_2 + 2$ and $g_3(x_3, y_3) = (x_3 - 3)^2 + (y_3 - 3)^2 - 1$ and we are given three disjoint curves, C_1, C_2, and C_3 given by $g_1(x_1, y_1) = 0$, $g_2(x_2, y_2) = 0$ and $g_3(x_3, y_3) = 0$ respectively. We would like to find the shortest total distances from C_1 to C_2 and C_1 to C_3 in the closed and bounded set of $[-2, 5] \times [-1, 4]$. We show the curves C_1, C_2, and C_3 in **Figure 6** below.*

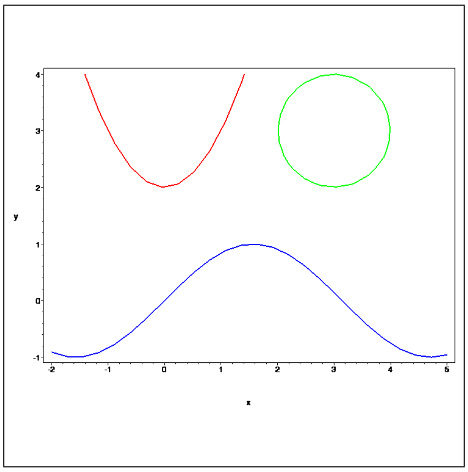

Figure 6. Three non-intersecting curves C_1, C_2, and C_3

Our objective is to minimize

$$\begin{aligned} f(x_1, y_1, x_2, y_2, x_3, y_3) &= (x_1 - x_2)^2 + (x_1 - x_3)^2 + (y_1 - y_2)^2 + (y_1 - y_3)^2 \quad (37) \\ \text{subject to } g_1(x_1, y_1) &= 0, g_2(x_2, y_2) = 0 \text{ and } g_3(x_3, y_3) = 0. \quad (38) \end{aligned}$$

We write $\mathbf{x}_i = (x_i, y_i), i = 1, 2, 3$. It is easy to verify that $\sum_{i=1}^{3} \frac{\partial}{\partial \mathbf{x}_i} (f(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)) = 0$, and it follows from Corollary 5 that $\sum_{i=1}^{3} \lambda_i \frac{\partial}{\partial \mathbf{x}_i} g_i(\mathbf{x}_i) = 0$. Moreover, we have $\sum_{j=1}^{3} \lambda_j \nabla g_j(\mathbf{x}_0) = 0$ if $\mathbf{x}_0 = (x_1, y_1, x_2, y_2, x_3, y_3)$ is an extremum of f. Consequently, we may write

$$\nabla g_1(\mathbf{x}_0) = -\frac{\lambda_2}{\lambda_1} \nabla g_2(\mathbf{x}_0) - \frac{\lambda_3}{\lambda_1} \nabla g_3(\mathbf{x}_0), \quad (39)$$

when $\lambda_1 \neq 0$. This equation says that the minimum occurs when the normal vector for C_1 at x_0 is a linear combination of the the normal vectors for C_2 and C_3 at x_0. We shall demonstrate this geometrically. First, we show the result we obtained from Maple (see [9]). We obtain the shortest total distances occurs when

$$\begin{aligned} A &= \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} 1.503078740 \\ .9977080403 \end{bmatrix} \in C_1, \\ B &= \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} .4425626436 \\ 2.195861693 \end{bmatrix} \in C_2, \\ C &= \begin{bmatrix} x_3 \\ y_3 \end{bmatrix} = \begin{bmatrix} 2.401228926 \\ 2.199079779 \end{bmatrix} \in C_3, \text{ and} \\ \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix} &= \begin{bmatrix} -4.799050783 \\ 2.396307306 \\ 1.499989270 \end{bmatrix}. \end{aligned}$$

It follows from our computation that indeed $\lambda_1 (\nabla g_1 |_{\text{at } A}) = \lambda_2 \overrightarrow{AB} + \lambda_3 \overrightarrow{AC}$, which can be shown below.

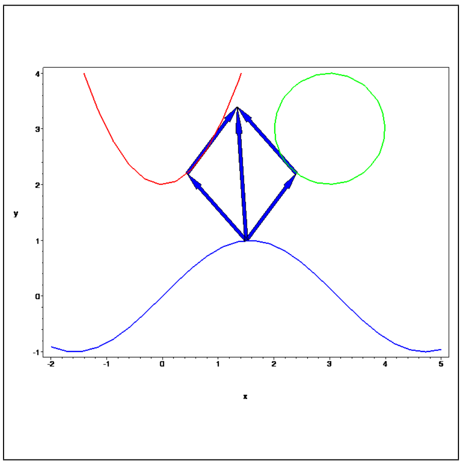

Figure 7. The shortest total squared distances among three curves

Here we demonstrate the use of Corollary 5 in 3-D.

Example 6 *Let* $g_1(x,y,z) = x^2+y^2+z^2-1$, $g_2(x,y,z) = x^2+(y-3)^2+(z-1)^2-1$, $g_3(x,y,z) = z-(x^2+y^2)-2$ *and* $g_4(x,y,z) = (4(x-3)+(y-3)+(z-1))(x-3)+((x-3)+4(y-3)+(z-1))(y-3)+((x-3)+(y-3)+4(z-1))(z-1)-3$. *We are given four disjoint convex surfaces,* S_1, S_2, S_3 *and* S_4 *given by* $g_1(x,y,z) = 0$, $g_2(x,y,z) = 0$, $g_3(x,y,z) = 0$ *and* $g_4(x,y,z) = 0$ *respectively. We would like to find* ***the shortest total squared distances*** *from* S_1 *to* S_2, S_1 *to* S_3,*and* S_1 *to* S_4 *in the closed and bounded domain of* $[-3,4]\times[-2,5]\times[-3,5]$. *We show the surfaces* S_1, S_2, S_3, *and* S_4 *in* ***Figure 8***.

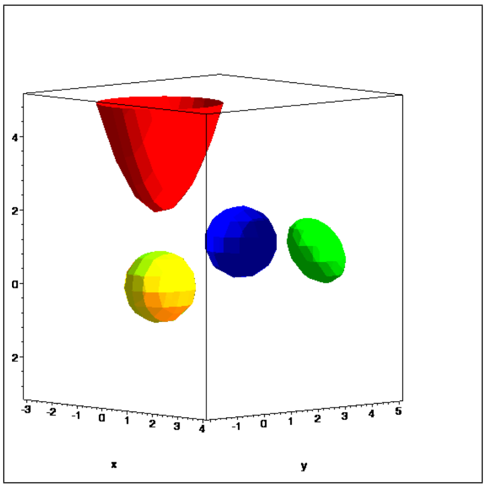

Figure 8. Four convex surfaces

If we use the notation of $\mathbf{x}_1 = (x_1^1, x_2^1, x_3^1)$, $\mathbf{x}_2 = (x_1^2, x_2^2, x_3^2)$, $\mathbf{x}_3 = (x_1^3, x_2^3, x_3^3)$, and $\mathbf{x}_4 = (x_1^4, x_2^4, x_3^4)$, we want to find the minimum of

$$f(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3,\mathbf{x}_4) = |\mathbf{x}_1-\mathbf{x}_2|^2 + |\mathbf{x}_1-\mathbf{x}_3|^2 + |\mathbf{x}_1-\mathbf{x}_4|^2 \tag{40}$$

subject to $g_1(\mathbf{x}_1) = 0$, $g_2(\mathbf{x}_2) = 0$, $g_3(\mathbf{x}_3) = 0$ and $g_4(\mathbf{x}_4) = 0$. It is easy to verify (with the help of a CAS-Maple) that $\sum_{i=1}^4 \frac{\partial}{\partial \mathbf{x}_i}(f(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3,\mathbf{x}_4)) = 0$ and we have $\sum_{j=1}^4 \lambda_j \nabla g_j(\mathbf{x}_0) = 0$ for some $\mathbf{x}_0$ in R^3. It follows from Corollary 5 that if $\lambda_1 \neq 0$, we may write

$$\nabla g_1(\mathbf{x}_0) = -\frac{\lambda_2}{\lambda_1}\nabla g_2(\mathbf{x}_0) - \frac{\lambda_3}{\lambda_1}\nabla g_3(\mathbf{x}_0) - \frac{\lambda_4}{\lambda_1}\nabla g_3(\mathbf{x}_0). \tag{41}$$

We use Maple for computation and we obtain the shortest total squared distances occurs when

$$\mathbf{x}_1 = \begin{bmatrix} .3966581137 \\ .7009795128 \\ .5926972781 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} .1674881704 \\ 2.029242761 \\ .8280171630 \end{bmatrix}$$

$$\mathbf{x}_3 = \begin{bmatrix} .1017095600 \\ .1797424919 \\ 2.042652198 \end{bmatrix}, \mathbf{x}_4 = \begin{bmatrix} 2.387696012 \\ 2.486312830 \\ 1.099333105 \end{bmatrix}, \text{ and}$$

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{bmatrix} = \begin{bmatrix} 3.698195879 \\ 1.368275399 \\ -2.899909840 \\ .6952991848 \end{bmatrix}. \tag{42}$$

We show the identify (41) geometrically below:

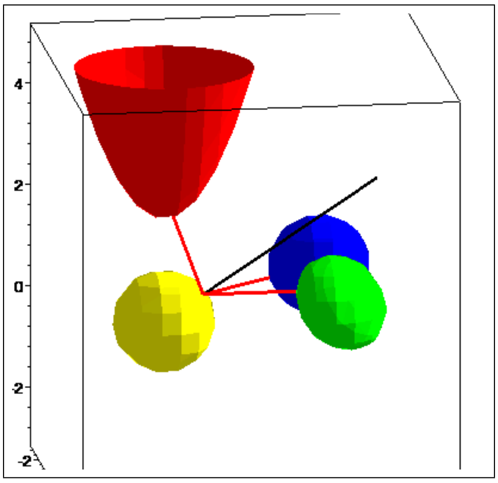

Figure 9. The shortest total squared distances among four convex surfaces

Remark: It is understood that Corollary 5 gives only a necessary condition when optimizing a function. For the previous example, we may obtain another set of solutions as follows:

$$\begin{aligned}
\mathbf{x}_1 &= \begin{bmatrix} -.3189240652 \\ -.7776125973 \\ -.5418543061 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} .07792698416 \\ 3.923034632 \\ 1.376741893 \end{bmatrix} \\
\mathbf{x}_3 &= \begin{bmatrix} -.05209975982 \\ -.1270315852 \\ 2.018851409 \end{bmatrix}, \mathbf{x}_4 = \begin{bmatrix} 2.521791552 \\ 2.415696446 \\ .9328295271 \end{bmatrix}, \text{ and} \\
\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{bmatrix} &= \begin{bmatrix} -10.98816726 \\ -5.092601153 \\ -5.121411430 \\ 1.107790413 \end{bmatrix}.
\end{aligned} \tag{43}$$

This clearly does not give the shortest square distance, which can be seen from the figure below, where the point $\mathbf{x}_1$ starts from the back of S_1 (yellow ball).

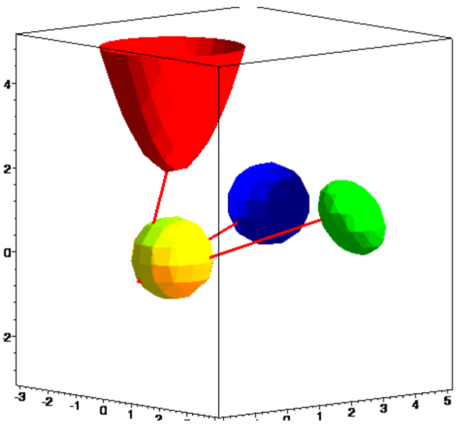

Figure 10. The largest total squared distances among four convex surfaces

5.1 Discussion

The Corollary 5 describes a nice geometric interpretation between the global values of the total squared distances function f and the gradient vectors at those critical points satisfying side conditions. When we replace f by an arbitrary function in Theorem 4, we can only focus on the relationship between equations (27) and (28).

We demonstrate this by using the following example; we use $\mathbf{x}_1 = (x_1^1, x_2^1)$, $\mathbf{x}_2 = (x_1^2, x_2^2)$ and $\mathbf{x}_3 = (x_1^3, x_2^3)$ to denote points in $\mathbb{R}^2$. In the next example, we want to minimize a general function $f(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ subject to $g_1(\mathbf{x}_1) = 0, g_2(\mathbf{x}_2) = 0$ and $g_3(\mathbf{x}_3) = 0$, which represent three curves in $\mathbb{R}^2$.

Example 7 *Let $g_1(x_1^1, x_2^1) = x_2^1 + (x_1^1)^2$, $g_2(x_1^2, x_2^2) = (x_1^2)^2 - x_2^2 + 2$ and $g_3(x_1^3, x_2^3) = (x_1^3 - 3)^2 + (x_2^3 - 3)^2 - 1$ and we are given three disjoint curves, $C_1, C_2,$ and C_3 given by $g_1(x_1^1, x_2^1) = 0, g_2(x_1^2, x_2^2) = 0$ and $g_3(x_1^3, x_2^3) = 0$ respectively. We first define*

$$f_1(x_1^1, x_2^1, x_1^2, x_2^2) = \left(x_1^2 - x_1^1\right)^3 + \left(x_2^2 - x_2^1\right)^2 \text{ and} \tag{44}$$

$$f_2(x_1^1, x_2^1, x_1^3, x_2^3) = \left(x_1^3 - x_1^1\right)^3 + \left(x_2^3 - x_2^1\right)^2 \tag{45}$$

We would like to minimize the function

$$f(x_1^1, x_2^1, x_1^2, x_2^2, x_1^1, x_2^1, x_1^3, x_2^3) = f_1(x_1^1, x_2^1, x_1^2, x_2^2)) + f_2(x_1^1, x_2^1, x_1^3, x_2^3) \tag{46}$$

subject to $g_1(x_1^1, x_2^1) = 0, g_2(x_1^2, x_2^2) = 0$ and $g_3(x_1^3, x_2^3) = 0$ in the closed and bounded set of $[-2, 4] \times [-2, 4]$.

We note that the function f is being twisted a bit from the total squared distances function and yet satisfying $\sum_{i=1}^{3} \frac{\partial}{\partial \mathbf{x}_i} \left(f(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)\right) = 0$. If $\mathbf{x}_0 = (\mathbf{x}_1^*, \mathbf{x}_2^*, \mathbf{x}_3^*)$ is a critical point for f, we do not expect $\nabla g_1(\mathbf{x}_0)$ to be written as a linear combination of $\nabla g_2(\mathbf{x}_0)$ and $\nabla g_3(\mathbf{x}_0)$ in this case, but we do have the equality of $\sum_{j=1}^{3} \lambda_j \frac{\partial}{\partial \mathbf{x}_i} \left(g_j(\mathbf{x}_j^*)\right) = 0$ as expected from Theorem 4. We use Maple for computation (see [11]) to obtain the following information:

$$\begin{aligned} \mathbf{x}_1^* &= \begin{bmatrix} .4815982352 \\ -.2319368601 \end{bmatrix}, \mathbf{x}_2^* = \begin{bmatrix} -.1213699858 \\ 2.014730673 \end{bmatrix}, \mathbf{x}_3^* = \begin{bmatrix} 2.157465703 \\ 2.461357300 \end{bmatrix}, \text{ and} \\ \lambda_1 &= 9.879923388, \lambda_2 = 4.493335067, \lambda_3 = 5.000149748, \text{ and } f(\mathbf{x}_0) = 16.78885390. \end{aligned} \tag{47}$$

We demonstrate the solution set $\{\mathbf{x}_1^*, \mathbf{x}_2^*, \mathbf{x}_3^*\}$ and its corresponding curves ($g_1(\mathbf{x}_1) = 0$ in blue,

$g_2(\mathbf{x}_2) = 0$ in red and $g_3(\mathbf{x}_3) = 0$ in green) by using the following **Figure 11**.

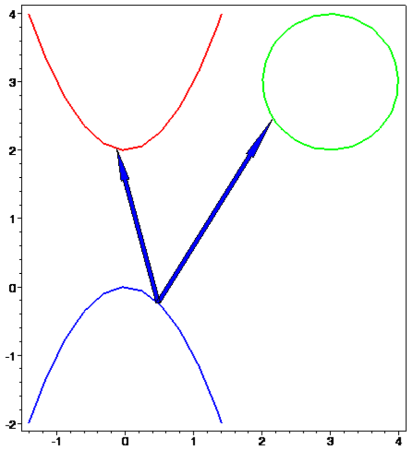

Figure 11. The objective function is not the total squared distances function

Finally, we use the following example to show that the objective function f satisfying the condition of Theorem 5 can be very arbitrary. We use the notation of $\mathbf{x}_1 = (x_1^1, x_2^1, x_3^1)$, $\mathbf{x}_2 = (x_1^2, x_2^2, x_3^2)$, $\mathbf{x}_3 = (x_1^3, x_2^3, x_3^3)$, and $\mathbf{x}_4 = (x_1^4, x_2^4, x_3^4)$. Let $g_1(\mathbf{x}_1)$, $g_2(\mathbf{x}_2)$, $g_3(\mathbf{x}_3)$, and $g_4(\mathbf{x}_4)$ be defined in Example 6. We will explore some but not all critical points for

$$f(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4) = \left(x_1^1 - x_2^1\right)^2 + (x_2^1 - x_3^1)^2 + (x_2^2 - x_3^2) - e^{2x_1^2 - 2x_2^2} \tag{48}$$

$$+ \left(x_2^3 - x_3^3\right)^2 + (x_1^4 - x_3^4) + 2e^{x_1^4 - x_2^4}, \tag{49}$$

subject to $g_1(\mathbf{x}_1) = 0$, $g_2(\mathbf{x}_2) = 0$, $g_3(\mathbf{x}_3) = 0$, and $g_4(\mathbf{x}_4) = 0$, in the closed and bounded domain of $[-3, 4] \times [-2, 5] \times [-3, 5]$ for $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$, and $\mathbf{x}_4$. Notice that the function f is picked arbitrarily so the condition (27) is satisfied, which can be verified by Maple easily. We further note that the side condition of $g_1(\mathbf{x}_1) = 0$ representing a sphere centered at origin and with radius 1. By restricting the domain of x_1^1 to be $[-1, 1]$, and with the help of Maple, we show the following cases. Due to complexity of the problem and possible computation limitations from Maple, we shown only the following critical points (see [12]).

Case 1. When

$$\mathbf{x}_1^* = \begin{bmatrix} 0.7071067812 \\ 0 \\ -0.7071067812 \end{bmatrix}, \mathbf{x}_2^* = \begin{bmatrix} 0.01491365605 \\ 2.285554355 \\ 1.699531989 \end{bmatrix}$$

$$\mathbf{x}_3^* = \begin{bmatrix} 0 \\ 0.5 \\ 2.25 \end{bmatrix}, \mathbf{x}_4^* = \begin{bmatrix} 3.744185903 \\ 2.336976634 \\ 0.9188374623 \end{bmatrix},$$

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{bmatrix} = \begin{bmatrix} -1 \\ 0.7147635962 \\ -3.5 \\ -2.053492552 \end{bmatrix}, \text{ and } f(\mathbf{x}_1^*, \mathbf{x}_2^*, \mathbf{x}_3^*, \mathbf{x}_4^*) = 15.63229233. \tag{50}$$

We note that $\sum_{j=1}^{4} \lambda_j \dfrac{\partial}{\partial \mathbf{x}_i} \left(g_j(\mathbf{x}_j^*)\right) = 0$ in this case; however, as expected for $\mathbf{x}_0 = (\mathbf{x}_1^*, \mathbf{x}_2^*, \mathbf{x}_3^*, \mathbf{x}_4^*)$,since f is not the distance function, $\sum_{j=1}^{4} \lambda_j \nabla g_j(\mathbf{x}_0) \neq \mathbf{0}$, the zero vector in R^3.

Case 2. When

$$\begin{aligned}
\mathbf{x}_1^* &= \begin{bmatrix} .4082482905 \\ -.8164965809 \\ .4082482905 \end{bmatrix}, \mathbf{x}_2^* = \begin{bmatrix} -.0008495622453 \\ 3.707531180 \\ .2933183827 \end{bmatrix} \\
\mathbf{x}_3^* &= \begin{bmatrix} 0 \\ 0.5 \\ 2.25 \end{bmatrix}, \mathbf{x}_4^* = \begin{bmatrix} 2.189685551 \\ 3.318307528 \\ 1.492006921 \end{bmatrix}, \\
\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{bmatrix} &= \begin{bmatrix} -3 \\ -.7075322009 \\ -3.5 \\ .3387486223 \end{bmatrix}, \text{ and } f(\mathbf{x}_1^*, \mathbf{x}_2^*, \mathbf{x}_3^*, \mathbf{x}_4^*) = 10.82074775. \qquad (51)
\end{aligned}$$

Case 3. The following gives us only a critical point for f.

$$\begin{aligned}
\mathbf{x}_1^* &= \begin{bmatrix} .7071067812 \\ 0 \\ -.7071067812 \end{bmatrix}, \mathbf{x}_2^* = \begin{bmatrix} -.0008495622453 \\ 3.707531180 \\ .2933183827 \end{bmatrix} \\
\mathbf{x}_3^* &= \begin{bmatrix} 0 \\ 0.5 \\ 2.25 \end{bmatrix}, \mathbf{x}_4^* = \begin{bmatrix} 2.189685551 \\ 3.318307528 \\ 1.492006921 \end{bmatrix}, \\
\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{bmatrix} &= \begin{bmatrix} -1 \\ -.7075322009 \\ -3.5 \\ .3387486223 \end{bmatrix}, \text{ and } f(\mathbf{x}_1^*, \mathbf{x}_2^*, \mathbf{x}_3^*, \mathbf{x}_4^*) = 8.8207477557572. \qquad (52)
\end{aligned}$$

Case 4. The following gives us another critical point for f.

$$\begin{aligned}
\mathbf{x}_1^* &= \begin{bmatrix} .4082482905 \\ -.8164965809 \\ .4082482905 \end{bmatrix}, \mathbf{x}_2^* = \begin{bmatrix} .01491365605 \\ 2.285554355 \\ 1.699531989 \end{bmatrix} \\
\mathbf{x}_3^* &= \begin{bmatrix} 0 \\ 0.5 \\ 2.25 \end{bmatrix}, \mathbf{x}_4^* = \begin{bmatrix} 2.189685551 \\ 3.318307528 \\ 1.492006921 \end{bmatrix}, \\
\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{bmatrix} &= \begin{bmatrix} -3 \\ .7075322009 \\ -3.5 \\ .3387486223 \end{bmatrix}, \text{ and } f(\mathbf{x}_1^*, \mathbf{x}_2^*, \mathbf{x}_3^*, \mathbf{x}_4^*) = 7.982498677. \qquad (53)
\end{aligned}$$

In summary, we may replace the distance function from a surface to another one by a curve (in an appropriate dimension) connecting two surfaces instead of a straight line. Author believes that there will be some applications regarding the use of Theorem 4.

6 Conclusion

It is known that finding minimum distance between two surfaces has many applications in robotic engineering (see [3]). Traditionally when technological tools are not available, students may find applying Lagrange Multiplier Method in solving optimization problems difficult; not only due to complicated algebraic manipulation nature but also they often do not fully understand the geometric interpretation behind the method. This paper demonstrates that optimization problems can be made interesting if teachers inspire students with geometric motivation. Moreover, a computer algebra system allows us to concentrate on the main mathematics concepts on one hand and carry out the complicated computations in the mean time. This paper shows examples how we can help students to integrate geometric intuitions, concepts learned in Linear Algebra, with Lagrange Multipliers Method learned in Multi-variable Calculus.

References

[1] Yang, W.-C., *Some Optimization Problems in Two and Three Dimensions*, at the Proceedings of the Eleventh Asian Technology Conference in Mathematics, 2006, pp 299-307, ISBN 0-9763064-3-3.

[2] Gao, D. Y. & Yang, W.-C., '*Complete Solutions to Minimal Distance Problem between Two nonconvex Surfaces*', coauthored with David Y. Gao, at the Journal of Optimization by Taylor & Francis, September 2008.

[3] Patoglu, V. & Gillespie, R. B. '*Extremal Distance Maintenance for Parametric Curves and Surfaces'*, Proceedings of the 2002 IEEE International Conference on Robotics 8 Automation,Washington, DC May 2002, pp2817-2823.

Software Packages

[4] [ClassPad] ClassPad Manager, a product of CASIO Computer Ltd., http://www.classpad.net or http://www.classpad.org/.

[5] [Maple] Maple 11, a product of Maplesoft, http://www.maplesoft.com/.

Supplemental Electronic Materials

[6] Yang, W.-C, *A video clip that gives motivations in Section 2*.

[7] Yang, W.-C., *Maple worksheet* for Example 1.

[8] Yang, W.-C., *Maple worksheet* for Example 2.

[9] Yang, W.-C., *Maple worksheet* for Example 5.

[10] Yang, W.-C., *Maple worksheet* for Example 6.

[11] Yang, W.-C., *Maple worksheet* for Example 7.

[12] Yang, W.-C., *Maple worksheet* for this Example.