2022

Sensitivity Analysis of Basins of Attraction for Nelder-Mead

Sonia K. Shah
Bowdoin College

Sensitivity Analysis of Basins of Attraction for Nelder-Mead

An Honors Paper for the Department of Mathematics

Sonia K. Shah

Bowdoin College, 2022

Abstract

The Nelder-Mead optimization method is a numerical method used to find the minimum of an objective function in a multidimensional space. In this paper, we use this method to study functions - specifically functions with three-dimensional graphs - and create images of the basin of attraction of the function. Three different methods are used to create these images named the systematic point method, randomized centroid method, and systemized centroid method. This paper applies these methods to different functions. The first function has two minima with an equivalent function value. The second function has one global minimum and one local minimum. The last function studied has several minima of different function values. The systematic point method is a reliable method in particular scenarios but is extremely sensitive to changes in the initial simplex. The randomized centroid method was not found to be useful as the basin of attraction images are difficult to understand. This made it particularly troublesome to know when the method was working effectively and when it was not. The systemized centroid method appears to be the most precise and effective method at creating the basin of attraction in most cases. This method rarely fails to find a minimum and is particularly adept at finding global minima more effectively compared to local minima. It is important to remember that these conclusions are simply based off the results of the methods and functions studied and that more effective methods may exist.

Keywords: Optimization, Nelder-Mead, Basins of Attraction, Sensitivity Analysis, Minimization, Simplex

Acknowledgements

This paper would not have been possible without the support of so many others. First, I would like to thank my advisor, Professor Adam Levy, for his constant guidance and advice over the past year which have helped me grow as a mathematician and researcher. I would also like to thank my research companion, Gillian King, for her immeasurable support and help throughout this project. To my friends who listened to me talk and ramble about this paper and this topic - even when they had no clue what I was talking about - thank you. And finally, to my family - thank you for your endless love and support. I couldn't have done this without you.

Contents

List of Figures

Chapter 1

Introduction

The Nelder-Mead optimization method is a numerical method used to find the minimum of an objective function in a multidimensional space. In this paper, we use this method to study functions - specifically functions with three-dimensional graphs - and create images of the basins of attraction of the function. A basin of attraction is a collection of initializations from which these approximations approach a given attractor, or in this case, a minimum. Therefore, the basin of attraction images map out which minimum a particular initialization finds.

One constraint of using the Nelder-Mead method is that the method requires an initial simplex in order to find a minimum. There are many different ways to create the initial simplex used to generate the basin of attraction images. The first method I'll consider is called the systematic point method. It involved assigning two points of the simplex and cycling through the grid being studied to generate the third point. This third point is color-coded in the basin of attraction images to show which minimum is found. The second method studied is the randomized centroid method. The method randomly generates three points to create an initial simplex and then finds the centroid of the simplex which is color-coded to show which minimum was found. The last method we study is the systemized centroid method. In this case, the initial simplex is found by creating a fixed triangle around every integer point in the grid being studied. The centroid of that simplex is used to indicate which minimum is found in the basin of attraction images.

We use these three different methods of generating an initial simplex to study three different functions with different numbers and types of minima. The first function has two minima with an equivalent function value. The second function also has two minima but one minimum is global and the other is local. The last function has several minima of different function values. Each of these functions can help us to understand the Nelder-Mead method and the basin of attraction images generated in different ways.

As will be shown in the following chapters, the first method, the systematic point method, is useful in particular scenarios. This method is reliable but is extremely sensitive to changes in the initial simplex. Since two points of the simplex are assigned, any changes in those values can drastically shift the basin of attraction image. Therefore, this method can be considered unstable if the user is not cautious with the starting conditions and doesn't test multiple simplices. In the first function studied, the method was effective at finding both minima. However, in the second and third functions, the method often struggles to find the global minima over the local minima as we would want it to.

The second method, the randomized centroid method, is unfortunately not as effective as the systematic point method. Since the simplex of this method is created by randomly generated points and only the centroid of the simplex is color-coded, it is extremely difficult to use this method to better understand a function. The inability to know the initial simplex means we cannot tell when the method is working effectively and when it is not. Therefore, I did not end up using this method to study each function as it would not have provided new insight into the focus of this paper.

The systemized centroid method appears to be the most precise and effective method at creating the basin of attraction in most cases. We will see that this method rarely fails to find a minimum. Also, when the fixed simplex is smaller in size, the structure of the basin of attractions are closely related to the contour plots of the function. When we use this method to study the second and third functions we also see that it is adept at finding the global minima more effectively compared to the local minima than any other method used in this paper. In most cases, this method seems to create the basin of attraction images we would expect to see and find the minima that we would want it to find.

Chapter 2

Basins of Attraction

2.1 What are Basins of Attraction?

Basins of attraction are studied to better understand dynamical systems, especially differential equations. Differential equations are dynamical systems where time is measured continuously. To fully understand what a basin of attraction is, we must start by understanding what an attractor is. An attractor is a vector approached by approximations generated from a particular method or objective collective. A dynamical systems may have several attractors. For the purpose of this paper, we define an attractor to be the minimum points within a function. The objective in this paper is to run the Nelder-Mead method with several different starting points, or as we'll discover simplicies, in order to find minimums and be able to create basins of attraction images.

A basin of attraction is a collection of initializations from which these approximations approach a given attractor, or in this case, a minimum. This means that the qualitative behavior of a system can be different depending on which basin of attraction the initial condition lies in. Over the past year, my work has been leading towards generating images of basins of attraction using the Nelder-Mead optimization method.

Figure 2.1 is slightly different from the images that will be shown in the following sections of the paper but helpful to understand before continuing further. This image shows a basin of attraction for a differential equation where we see that the red and blue points are the equilibrium points of the function. The blue arrows show how an initialization moves over time and whether they move towards the equilibrium points or not. Throughout the rest of the paper, I use this same imaging idea but in terms of optimization. Therefore, instead of finding how initialization move with respect to equilibrium points over time, I looked at how a particular initialization found the minima of a function.

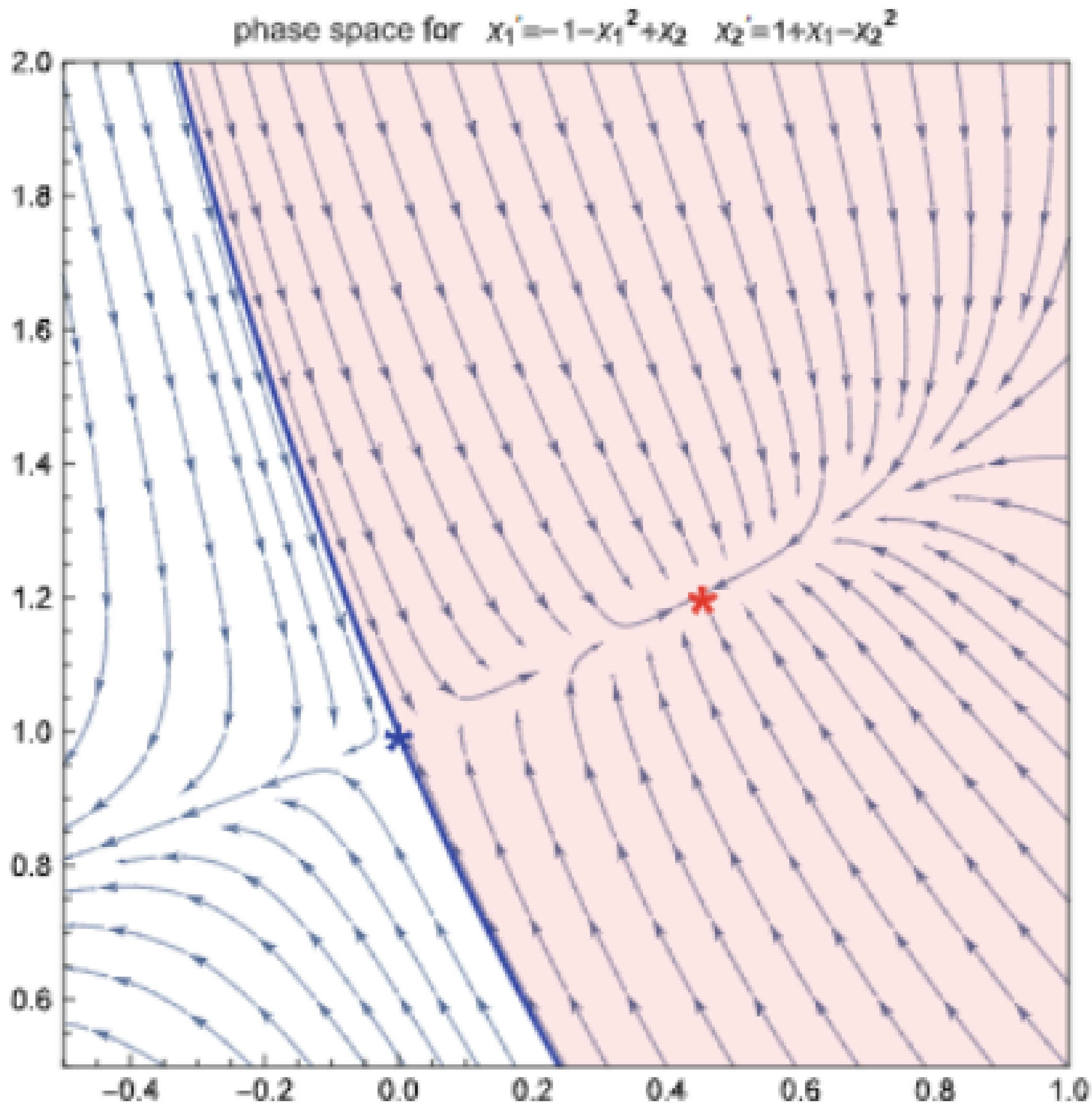


Figure 2.1: Basins of Attraction for a Differential Equation.

Chapter 3

Nelder-Mead Optimization Method

3.1 Nelder-Mead Method Overview

The Nelder-Mead method is a numerical method used to find the minimum of an objective function in a multidimensional space. In this paper, we restrict our focus to finding minima of three-dimensional functions. The Nelder-Mead method is a type of simplex method, and in this paper the simplex always has three vertices. It is also a direct search method meaning that it doesn't require gradients as other optimization methods do. Therefore, it is extremely useful for problems where the derivative can be hard to calculate.

One issue that is particularly troubling - and that we will study further in the paper - is that Nelder-Mead is a heuristic search method and can therefore converge to non-stationary, or non-minimum, points in certain cases. In addition, we cannot guarantee that the Nelder-Mead method will always converge to the global minimum of a function if a local minimum is present as well. In this paper, we mainly studied functions where we already knew the minima - both global and local - to avoid issues that may arise from this feature and to better understand our sensitivity tests.

To describe the Nelder-Mead method at a very high level, we first start with a triangle, or three points, placed onto the surface of a function. In other words, three point are chosen from the function to form a triangle with the respective function values of those points. We then study the function values of those trial points and work towards improving the worst vertex, or the vertex with the highest value, by lowering the function value at each step. We will discuss this method in further detail in the next few sections.

3.2 Outlining the Method

To start the Nelder-Mead method, we start by creating the simplex, or the original triangle, that we will be further manipulating. The starting triangle is often chosen to be relatively large as extremely small triangles tend to stall, fail, or converge to a non-stationary point. To create this initial triangle, I used several different methods which I will discuss in further sections.

After creating the initial simplex, the Nelder-Mead method begins to follow a pattern which I will outline here:

1. Order vertices of the simplex
2. Calculate the center of the best side
3. Run testing to improve function value of each point in the simplex
4. Test to terminate. If testing fails, begin from step 1

Every iteration of the method will begin with the first step of this process until it passes the termination testing in the last step. The following sections will discuss these steps in more detail.

3.3 Step 1: Order Vertices of the Simplex

After creating the initial simplex, the vertices of the simplex are ordered from best to worst based on the function value at their locations. When searching for minima, the lowest function value is considered the best and the highest function value is considered the worst. This step is necessary so that we know which vertex has the highest function value and we can work toward improving it to a lower function value in the following tests.

3.4 Step 2: Calculate the Center of the Best Side

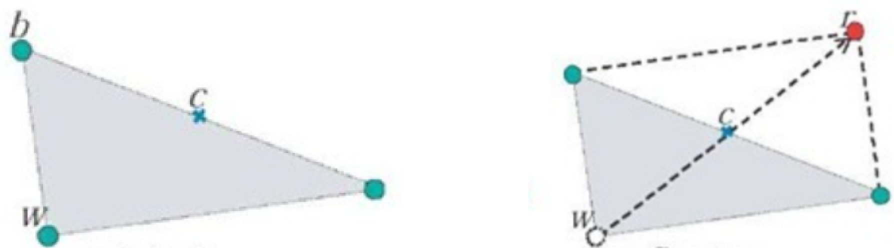
After ordering the vertices, we must calculate the center of best side of the triangle. The best side of the simplex is the side of the triangle built with the best and second-best points. We take the average of the x-coordinates and the average of the y-coordinates in order to find the center. For example, if the two best points were (1,1) and (3,3), the center of the best side would be (2,2). This center is necessary to compute as it will be used in the following testing.

3.5 Step 3: Testing to Improve Function Values of Simplex

Reflection and Replacement

The first test carried out within the Nelder-Mead method is called the reflection step. This process is seen in Figure 3.1. We reflect the worst vertex along the ray through the opposite edge in an attempt to find a new point, called the reflected point, with a lower function value than any existing point. We move in this direction as we are looking at a point in the opposite direction of the point with highest function value. Therefore, we would expect to find a lower value in this direction because the other points with a lower function value are also located in this direction.

On the left hand side, we see the original simplex of the function. The vertex with the highest function value in the simplex is labeled w, the vertex with the lowest function value in the simplex is labeled b, and the previously mentioned center of the best side is labeled



(a) Original Simplex (b) Reflected Point

Figure 3.1: Finding the Reflected Point

c. The unlabeled vertex of the triangle is the second best point and will be referred to as "point sb" in this paper. To create the image in (b), we use point w and point c to calculate the reflected point, labeled r, using the following formula:

$$x_r = x_c + \alpha(x_c - x_w) \tag{3.1}$$

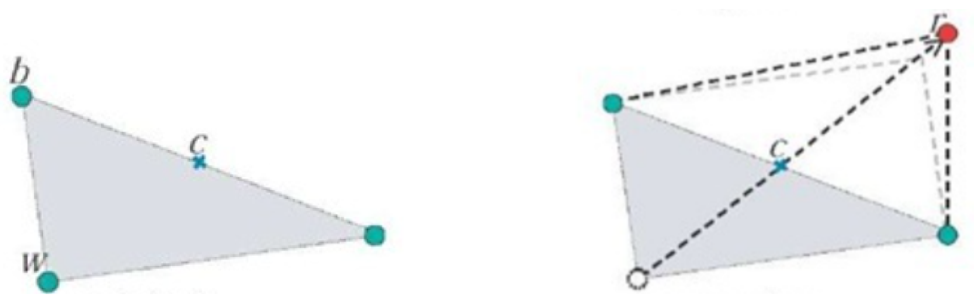
Note that in the above formula α is a reflection coefficient and has a typical value of 1.

Once we have the reflected point, we evaluate how to move forward. If $f(x_b) < f(x_r) < f(x_{sb})$, then x_r replaces x_w to form a new vertex. In other words, if the function value of the reflected point is lower than the function value of the second best vertex but higher than the function value of the best vertex, then we replace the worst vertex with the reflected point to create a new simplex. We can then start a new iteration of Nelder-Mead following the outline in Section 3.2.

If this is not the case, we move onto the next testing method. The testing method used is dependent on the function value of the reflected point.

3.5.1 Expansion

If the result of the reflection test was that $f(x_r) \leq f(x_b)$, then we attempt to expand past the reflected point. This process is illustrated in Figure 3.2b by the black lines, while the gray lines show where the previously calculated reflected point lies. While Figure 3.2 shows just a small step past the original reflected point, the expanded point usually stretches twice as far from the reflected edge of the triangle. The reasoning behind testing the expanded point is that if the reflected point produced progress towards the minimum of the function, then a step even further in that direction might also be an improvement.



(a) Original Simplex (b) Expanded Point

Figure 3.2: Finding the Expanded Point

In order to calculate the expanded point, x_e, we use the following formula:

$$x_e = x_c + \gamma(x_r - x_c) \tag{3.2}$$

where γ is the expansion coefficient and is any value greater than 1. I typically set $\gamma = 2$. In this formula, the expanded point is calculated by adding the center of the best side to the difference of the best side and the reflected point multiplied by the expansion coefficient.

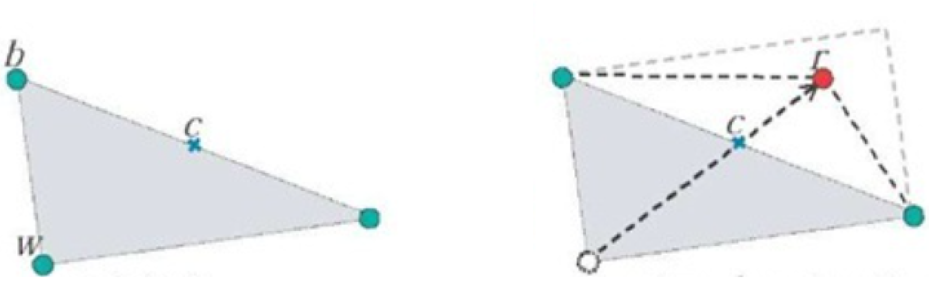
Since, we only run this test if the function value of the reflected point is better than the function value of the best point, there are only two possible cases:

$$f(x_e) < f(x_r) \;\; or \;\; f(x_e) \leq f(x_r)$$

If $f(x_e) < f(x_r)$, then we replace the vertex with the highest function value, x_w, with our expanded point and start a new iteration. If the latter case is true, then we simply replace the vertex with the highest function value with the reflected point and begin a new iteration.

3.5.2 Outer Contract

If the function value of the reflected point was not between the function values of the best and second-best point but rather between the function values of the worst point and the second best point, then we test the outer contract point. The outer contract point is usually half as far as the reflection edge of the triangle as illustrated by the black lines in Figure 3.3.



(a) Original Simplex (b) Outer Contract Point

Figure 3.3: Finding the Outer Contract Point

The rationale behind testing the outer contraction point is that the reflected point is not a strong enough improvement to keep and update our simplex with. However, we recognize that a smaller step in that direction may still be useful to find the minimum of the function. The outer contraction point, labeled x_{oc} is calculated using the following formula:

$$x_{oc} = x_c + \beta(x_r - x_c) \tag{3.3}$$

Where β is the contraction coefficient and $0 < \beta \leq 0.5$. I typically used a β value of 0.5. Here, the center of the best side is added to the difference of itself and the the reflected point multiplied by the contraction coefficient.

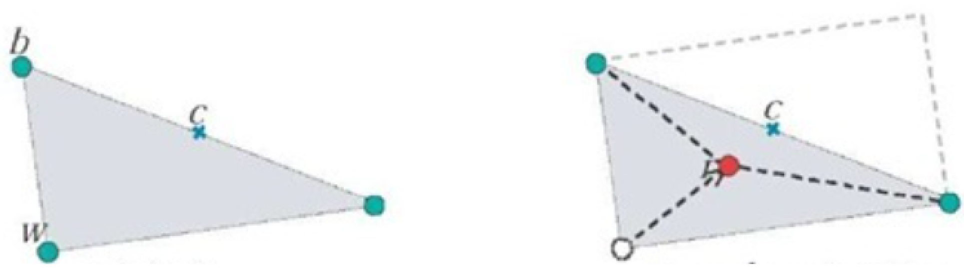
If the function value of the outer contract point is at least as good, if not better than the reflected point, then we update the vertex with the lowest function with the outer contract point and begin a new iteration.

$$\text{If: } f(x_{oc}) \leq f(x_r), \text{then: } x_w := x_{oc}$$

If the outer contract point is worse than the reflected point, we shrink the triangle, which will be discussed in section 3.5.4, before beginning a new iteration.

3.5.3 Interior Contract

If the function value of the reflected point was higher than the function value of the worst point in the function, $f(x_r) \geq f(x_w)$, then we test the interior contraction point. This point is usually halfway inside the reflection edge of the triangle and is depicted by the black lines in Figure 3.4b.



(a) Original Simplex (b) Interior Contract Point

Figure 3.4: Finding the Interior Contract Point

The rationale for this process is that the direction of the reflected point led to a result that was in actuality the reverse from the effect we were looking for. Therefore, we want to step away from that and we hope that a step in the opposite direction may lead us towards the minimum point. Another possible rationale is that the minimum point may already be inside our simplex and so we want to test points inside the simplex to see if they have a lower function value.

We use the following formula to find the interior contraction point:

$$x_{ic} = x_c + \beta(x_w - x_c) \tag{3.4}$$

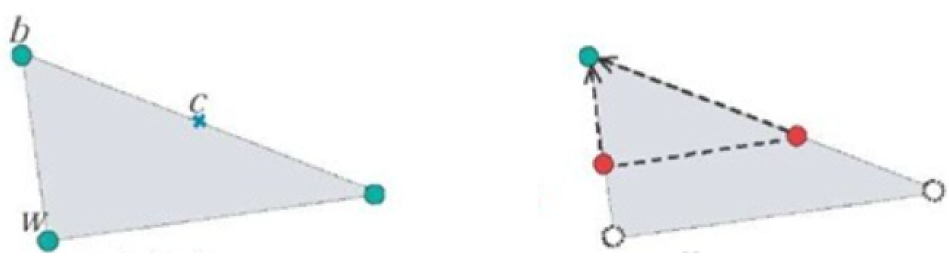
where β is the same contraction coefficient used to find the outer contraction point. In this case, we use almost the same formula that we did to find the outer contraction point, however, we replace the reflected point, x_r, with the worst point, x_w, in order to find a point inside of our simplex.

If the function value of the interior contraction point is better than the function value of the worst point, then we replace the worst point with the interior contraction point and begin a new iteration. Otherwise, we shrink the triangle and begin a new iteration.

$$\text{If: } f(x_{ic}) \leq f(x_w), \text{then: } x_w := x_{ic}$$

3.5.4 Shrink

If outer or interior contraction fail, then we shrink our simplex towards the current best vertex. To do this, we adjust two vertices, rather than just one. We often adjust the points with the two highest function values towards the point with the lowest function value in the simplex.



(a) Original Simplex (b) Shrink Point

Figure 3.5: Shrinking the Simplex

The reason that we carry this step out is because if outer and interior contraction both fail, we can suppose that the minimum point may be inside of our simplex. Therefore, shrinking towards the point with the lowest function value can help us find the minimum.

We use the following formula on the worst and second worst point, respectively:

$$x_{sp1} = x_b + \sigma(x_w - x_b) \tag{3.5}$$

$$x_{sp2} = x_b + \sigma(x_{sb} - x_b) \tag{3.6}$$

σ is the shrink coefficient and is typically set at 0.5. To find the shrink points, the best point is added to the difference between itself and the point we are looking to shrink multiplied by the coefficient.

Once all the points are replaced, except for the best point, we begin a new iteration.

3.6 Stopping Criteria

Since Nelder-Mead is an iterative method, the method will continue running until it meets a stopping criteria. Once we have narrowed down on a point that the method believes is a minimum, we need to be able to stop the method from running further.

There are multiple ways to do this, but the method I chose to use is by measuring how far the method is crawling over the function at each iteration. If the function was staying at the same spot over several iterations, I would stop the method from running. The reasoning for this is that if the function values were changing by an extremely small value, then that typically means the method has continuously been shrinking on a single point. This single point is most likely our minimum.

To do this, I would calculate the function value at the best point after every iteration and subtract it from the function value of the previous iteration. I accumulated a list of these differences. If the differences were below a stopping criteria for at least one hundred iterations then we stop the method from running.

To ensure that the stopping criteria didn't cause the method to stall early, I kept the stopping criteria at an extremely low value. Most of the time, the stopping criteria was set to a value of $1\text{x}10^{-300}$. In case the method stalled or never reached this stopping criteria, I also forced the method to stop after 10000 iterations.

Chapter 4

Creating the Initial Simplex

4.1 Overview

As mentioned above, I utilized several different methods to create the initial simplex to start the Nelder-Mead method. This allowed me to study how the basins of attraction for different functions are sensitive to changes in the starting simplex. I used three different methods to create initial simplices that I labelled as: systematic point, randomized centroid, and systemized centroid. However, as will be discussed, I mainly focused my further sensitivity testing on initial simplices generated by the systematic point and systemized centroid method.

4.2 Systematic Point Method

The first method I used to create a simplex was by having two assigned starting points and a systematically assigned integer third point. I usually changed the two assigned starting points depending on which function I was using, but, mostly, any points could be used as long as they didn't cause the optimization method to fail, stall, or converge to a non-stationary point. The third point was then assigned by running it through an iteration that ran through every integer value on the grid that I was looking at.

Assigning the third point in this way was extremely useful when creating images for basins of attraction. When running several different simplices together, I could color the systematically assigned point since every simplex shared the other two points. This made it easier to interpret and understand the images that I generate.

Initially, I simply used a random number generator to assign the third point. However, this was relatively inefficient and could potentially miss several points along the grid that I was studying. For example, if the assigned point somehow ended up in the same place often, the images created wouldn't give me as much detail as if the randomly assigned points was spread across the entire grid. Therefore, I switched over to choose the third point in a more methodical manner.

One issue that I did run into with the method that I used is that occasionally the simplex generated would fail to converge to a minimum. This could be occurring for various reasons. For example, if the systematic point was too close to the two assigned points, the simplex may be too small and could fail to converge to a minimum.

4.3 Randomized Centroid Method

Another method that I used to generate the initial simplex involved the centroid of the triangle. To do this, I ran a random integer generator to assign all three points of the initial simplex. I then plotted the centroids of each of these triangles and colored them based on the minimum found.

However, this method had many troubling issues. One concern similar to the systematic point method was if the assigned points all kept ending up in a similar spots, the images would not give me a lot of detail to understand sensitivity tests from. This was particularly concerning for this method as many of the centroids ended up towards the center of the grid as it is much harder to have a centroid towards the outskirts of the grid. This did not allow me a full picture of how this particular simplex found minima.

This process is illustrated in Figure 4.1. The image shows the randomized centroid method being used to create a basin of attraction for a function with a single minimum. Each point, blue or green, is a centroid of the initial simplex that has been plotted. As described above, we can see that the points are more densely gathered around the center of the range we are studying. As we move away from the center, the density of the points decreases greatly.

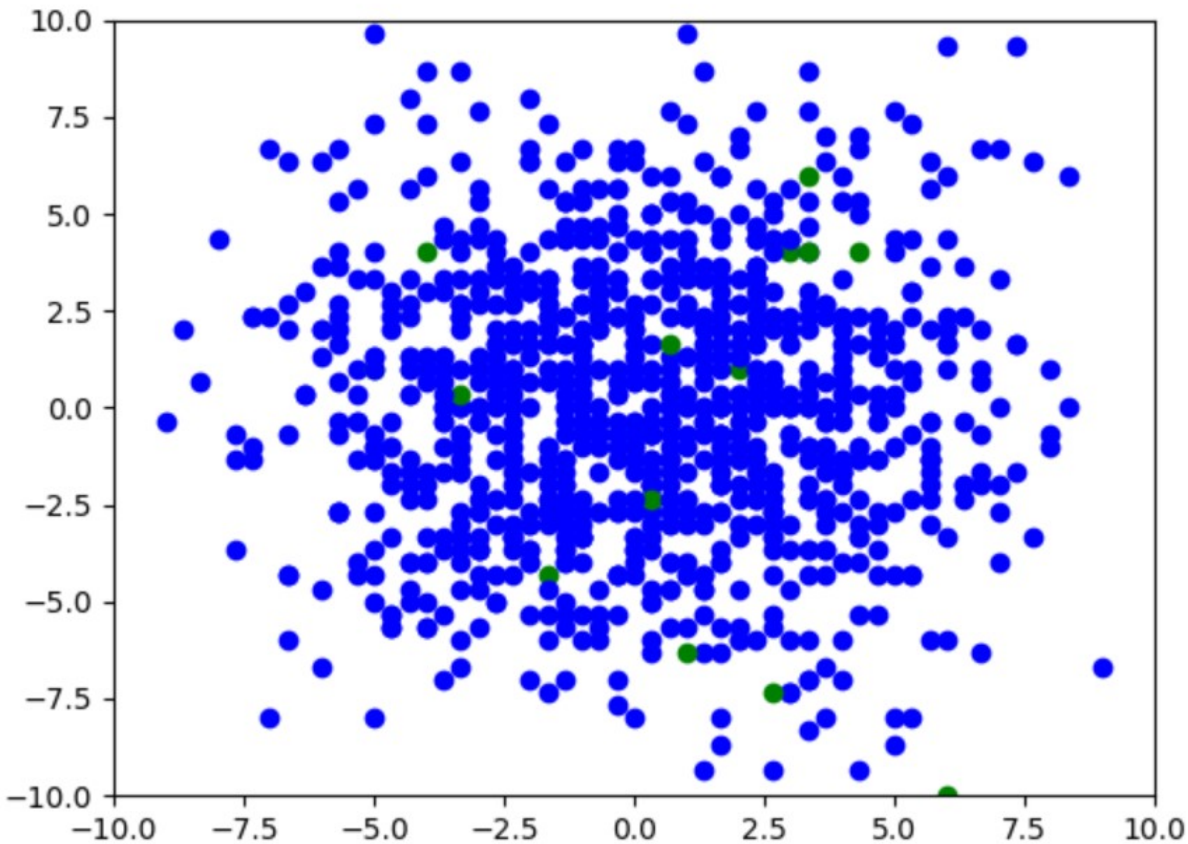


Figure 4.1: Randomized Centroid Test

As seen in Figure 4.1, this was also a very difficult method to understand from the pictures generated. From the basin images generated, it was impossible to know what the initial triangles looked like and therefore, impossible to decipher why certain centroids led to respective minima. For example, it was often the case that two points in extremely close proximity to each other found two different minima. Without knowing the simplex that led to both of those centroids, it was difficult to determine why that happened and run sensitivity tests using that method.

4.4 Systemized Centroid Method

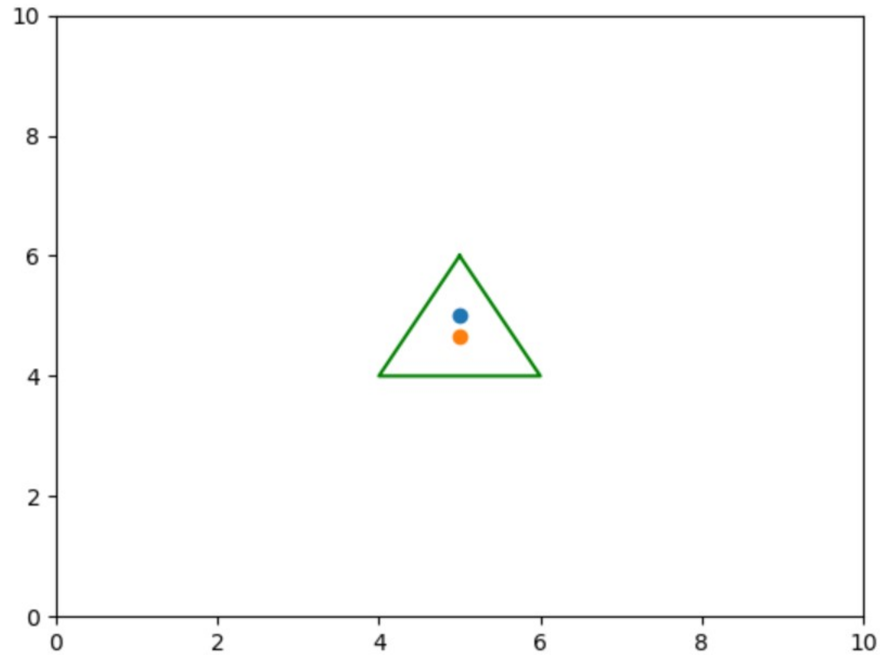
In order address unsystematic results that the random starting points entailed, I systemized the process. Instead of feeding the code three completely random points, I decided to build the simplices around given points in a structured fashion so the initial simplex would always be apparent.

I started by feeding the method points on the range that I wanted to look at and then built triangles of various sizes around it. Once I had a given point, I could create a simplex by choosing points directly above, below and to the left, and below and to the right of that point. I would add the value (0,1) to it to generate the vertex above. I would also add the values (-1, -1) and (1, -1) to generate the vertices below and to the left and below and to the right, respectively. For instance, if I started with a given point of (5,5), the final vertices for the simplex would be (5,6), (4,4), and (6,4).

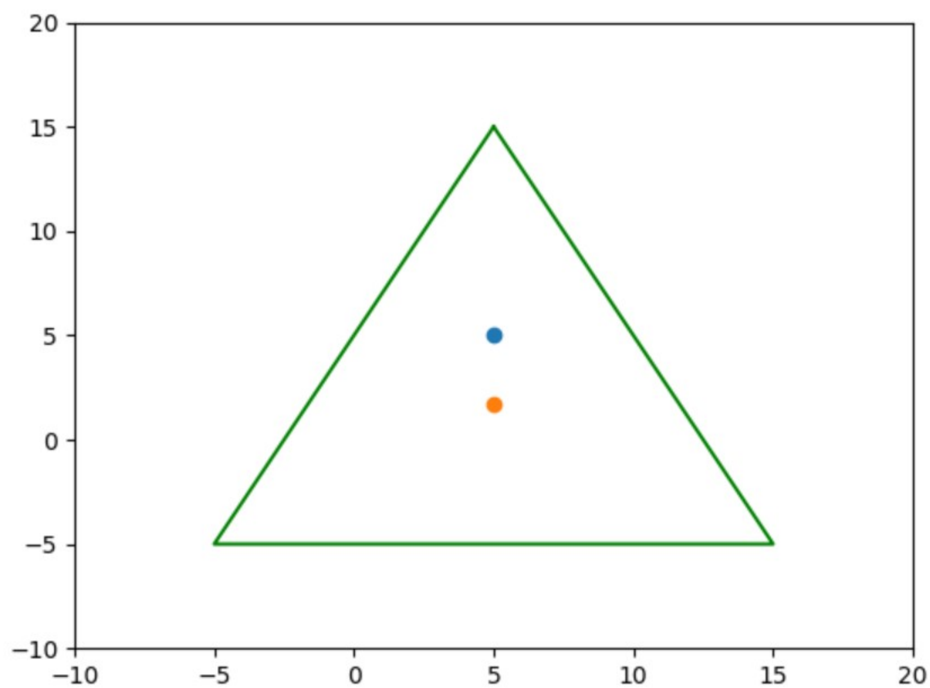
I could also manipulate the size of the initial simplex by multiplying the values added to the specific point by a constant k. For example, to generate a bigger simplex I could set k equal to 10 and then my values added to the given point would be (0,10), (-10,-10) and (10,-10), respectively. Once again, if I started with a given point of (5,5), my vertices for the simplex would be (5,15), (-5,-5), and (15,-5). This process is illustrated in Figure 4.2.

In Figure 4.2, the blue dot in both Figure 4.2a and Figure 4.2b represents the point (5,5) from which the simplex was built. In Figure 4.2a, k=1 and so the simplex, shown by the green triangle, has vertices of (5,6), (4,4), and (6,4). In Figure 4.2b, k=10 and so the simplex has vertices of (5,15), (-5,-5), and (15,-5). The orange dot in both images shows the actual centroid of the triangle that will be colored in further images.

To ensure that I didn't encounter the same issues that I did with the randomized centroid method, I chose the given point by feeding the method points on a grid. I could then move through the grid values and build simplices off of every integer point in the range that I was looking at. This ensured that points across the entire grid were incorporated and I was also given a much more informative picture of the results of how this particular way of generating a simplex was of finding minima.



(a) Initial simplex built around (5,5) with k = 1



(b) Initial simplex built around (5,5) with k = 10

Figure 4.2: Building a Simplex for the Systemized Centroid Method

Chapter 5

Mapping Minima and Tolerance Levels

5.1 Procedure for Mapping Minima

To create the basin of attraction images, I primarily worked with functions with two global minima. I also always knew the minima of my functions which made it easier to map them and create these images.

I started by running several different simplices through my code and then saving all of the minima found to a list. Then, I compared the minima found by my algorithm to the actual minima of the function I was working with. However, occasionally the points did not line up exactly to the actual minimum as they could be off by a few decimal points. Therefore, I allowed a small tolerance level to match real minimum points to those found by my code. Typically, I set this tolerance level to be 0.2. While this value may seem high, when considering the range of [-10,10] used for most basin of attraction images, it is quite small.

Based on the minimum that was found, I assigned the initial starting point a colored dot on the basin of attraction image. If the Nelder-Mead method failed - such as not coming within the tolerance range of a minimum or reaching one hundred thousand iterations - I used the a green dot to signify this. All other colors seen in the following images represent minima found by the method.

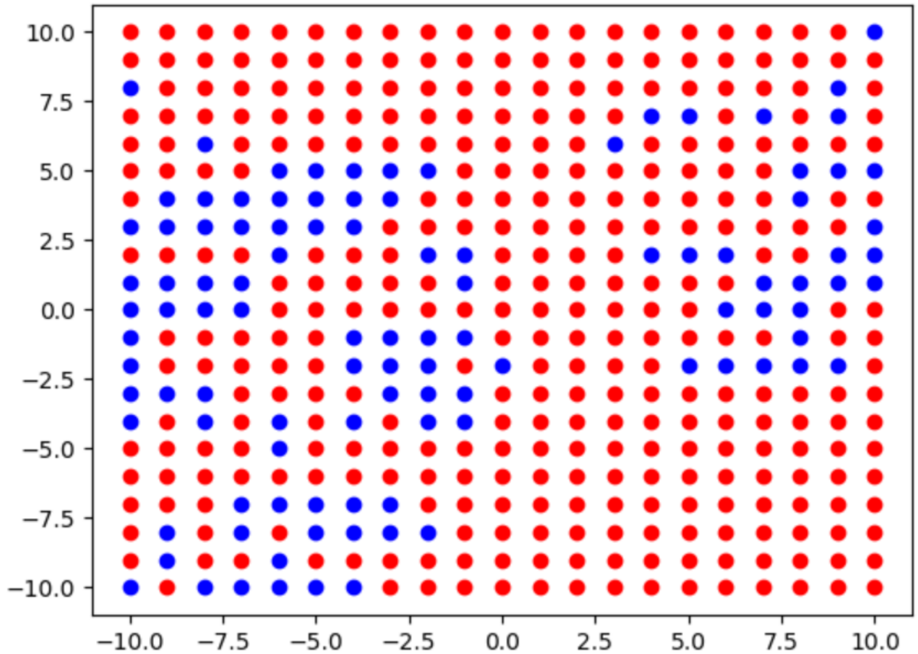
5.2 Sensitivity of Tolerance Levels

While I will discuss the basin of attraction images in further detail in the following sections, I first included some figures here to showcase how the basin of attraction images are sensitive to the tolerance level used. For example, in Figure 5.1A, the tolerance level was set at 0.9 which is considered large. This means that the point that Nelder-Mead converged on after stopping criteria were met was at most 0.9 units away from the real minimum. Since the tolerance level is so forgiving, we can see that the image results in all blue and red points. This means that for this particular function a minimum was always found with a large tolerance level.

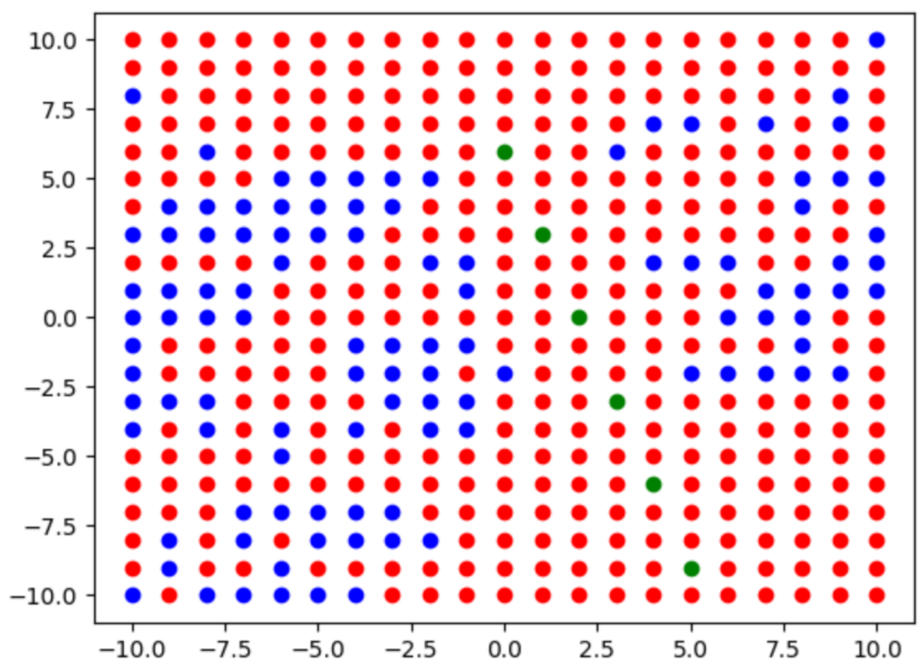
In Figure 5.1B, the tolerance level was set at 0.2. With a much tighter tolerance level, we see that the function has trouble converging to a minimum of the function from certain starting points - represented by the green dots. These points converged somewhere between 0.2 units and 0.9 units away from the actual minimum. By using a tighter tolerance level, we create a more accurate portrait of the basin of attraction.

Using a tighter tolerance level ensures that we won't color code a point that converges further away from the known minima of the function. For example, having a low tolerance helps guarantee that we don't accidentally include a point that was actually converging outside of our minimum to a point that just happened to be close to our minimum. It is also extremely useful for when we map functions where there are multiple minima that are really close to each other.

In the following basin of attraction images used within this paper, a tolerance level of 0.2 will always be used.



(a) Loose Tolerance Level: Tolerance = 0.9 units



(b) Tight Tolerance Level: Tolerance = 0.2

Figure 5.1: Sensitivity of Tolerance Levels

Chapter 6

Function with Two Similar Minima

6.1 Understanding the Function

The first function that I studied was:

$$f(x, y) = 2x^2 - 4xy + y^4 + 2$$

This was a good starting point as the minima of the function both have the same function value at the minima. The minima are located at (-1,-1) and (1,1) and have a function value of 1 at these points. Figure 6.1 shows a graph of this function. These images, along with the contour plots seen in Figure 6.2A and 6.2B help us to imagine what the basin of attraction images might look like, if we had a perfect algorithm.

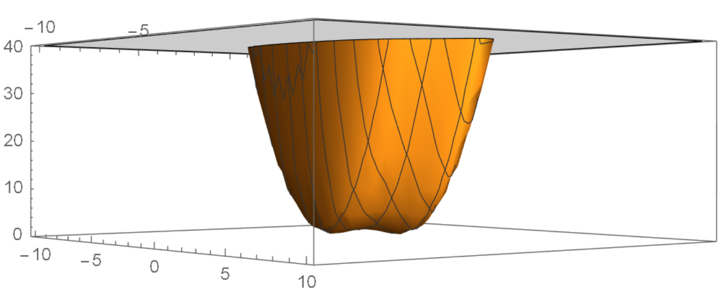
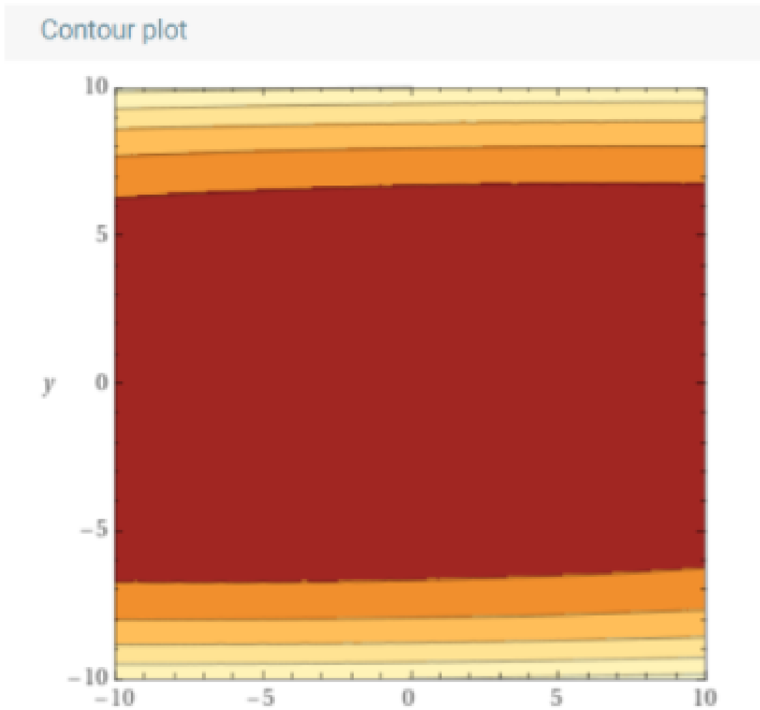
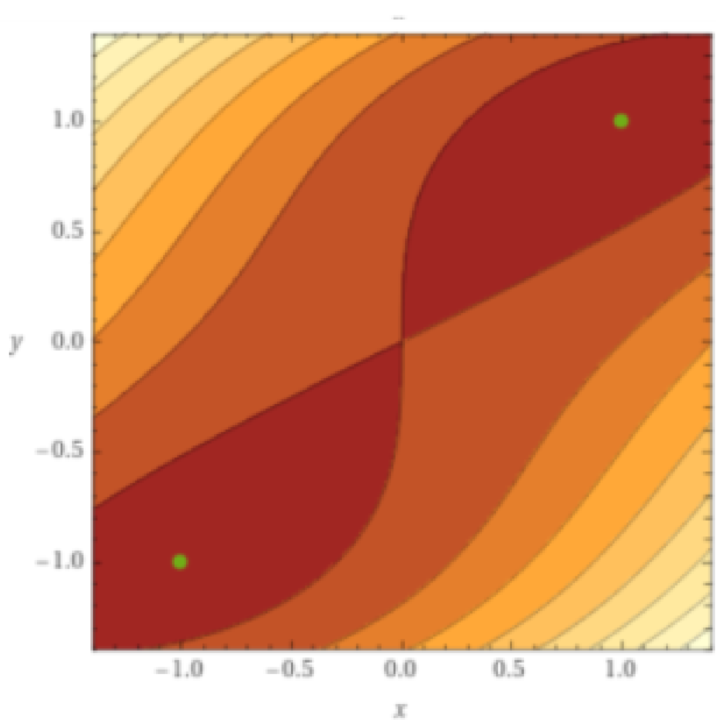


Figure 6.1: 3D Plot of $f(x, y) = 2x^2 - 4xy + y^4 + 2$



(a) Zoomed Out Contour Plot



(b) Zoomed in Contour Plot

Figure 6.2: Contour Plots of $f(x, y)$

Figure 6.2A shows the contour plot using axes with a range from $x = -10, 10$. This matches the same axes that will be used for basin images. From this zoomed out contour plot, we can see that there isn't a deep well or sink for the two minima. This is further reasoning for why the tight tolerance level is important. We want to ensure that our Nelder-Mead method is finding the minimum points even with a shallow well before we run further sensitivity tests on it.

Figure 6.2B shows a contour plot where the axes are zoomed in allowing us to see what the plot looks like near the minimum points. The minimum points are labeled with green dots. From these images, it appeared that the lower left triangle would be colored in a way that indicated the simplices went to the minimum at (-1,-1) and the upper right triangle would be colored in a way that indicated that they went to the minimum at (1,1).

6.2 Systematic Point Method

While using the systematic point method, we can generate many different images for this function based on the two assigned starting points used in the simplex. In the simplices for the following images, two points were assigned and the third point was pulled from the grid previously mentioned. The basin of attraction images will also be colored to show which minimum each simplex found. If the systemized point is colored blue, it means that the simplex found the minimum at (-1,-1). If the systematic point is colored red, the simplex found the minimum at (1,1). Lastly, if the systematic point is colored green, the simplex failed to find a minimum.

Figure 6.3 shows one of the images generated by using this method. In this image, I assigned fixed points of (0,0) and (2,-3) in the initial simplex. This image is extremely interesting as it does not follow what we assumed the basin of attraction would look like. Instead, the basin appears to be broken up into four different quadrants. The points in the top right and bottom left seem to have found the minimum located at (-1,-1), while the points in the top left and bottom right appear to have found the minimum at (1,1). Also, it appears that the simplex was unable to find any minima along the diagonal from the top left to the bottom right.

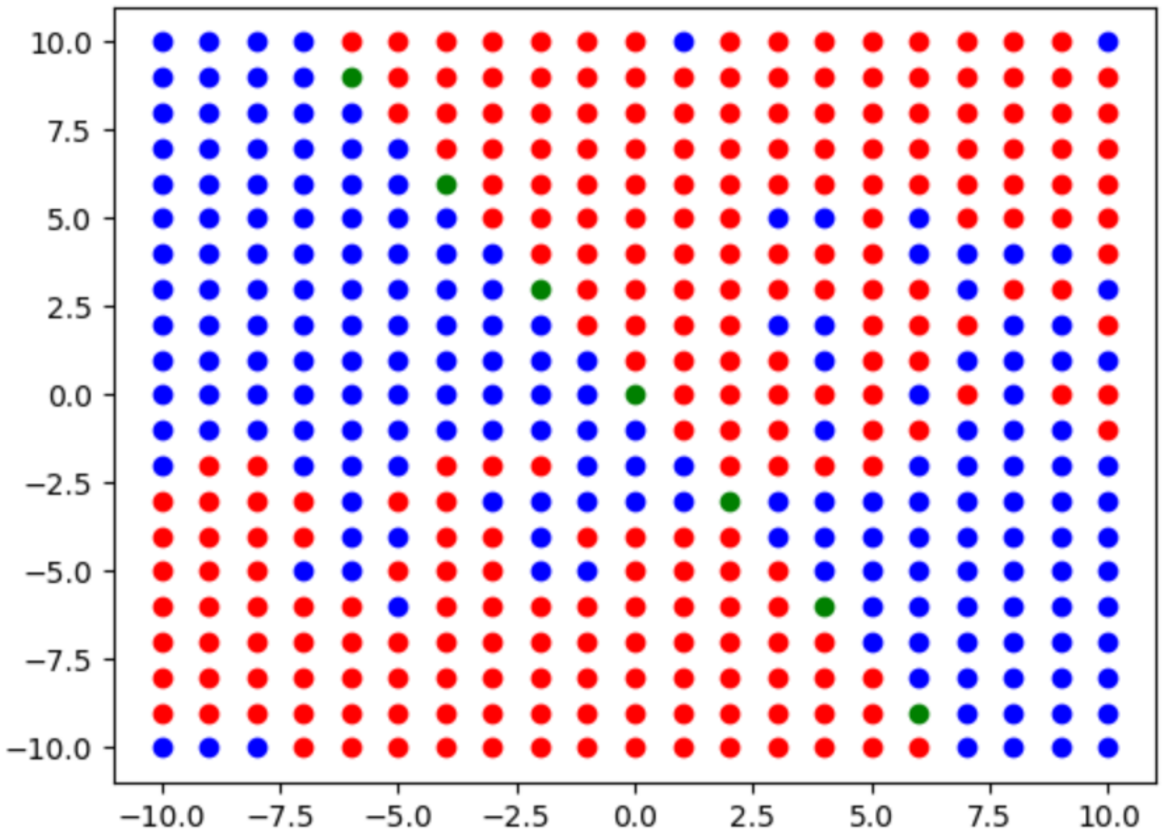


Figure 6.3: Systematic Point Method for $f(x, y) = 2x^2 - 4xy + y^4 + 2$ with assigned simplex points (0,0) and (2,-3)

I also created basins of attraction for this function with different assigned points creating the initial simplex. Figure 6.4 shows the basin of attraction for the function with the two assigned points being (0,0) and (2,3). Figure 6.5 shows the basin of attraction for the function with the two assigned points being (0,0) and (-2,-3).

We can see that both of these basins of attraction have many differences from the basin of attraction depicted in Figure 6.3. In Figure 6.4, the minimum (1,1) is found much more often whereas in Figure 6.5, (-1,-1) is found more often. Also, the green points, where the Nelder-Mead method was unable to find a minimum point, now stretches from the bottom

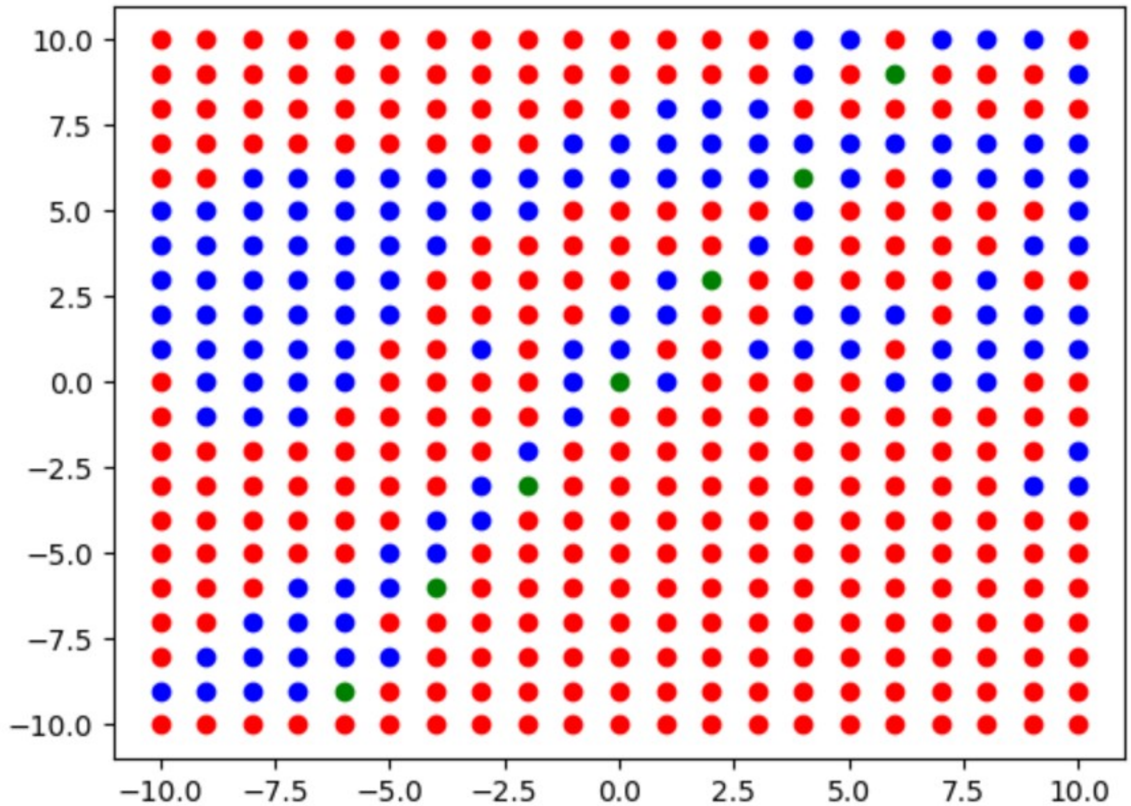


Figure 6.4: Systematic Point Method for $f(x, y) = 2x^2 - 4xy + y^4 + 2$ with assigned simplex points (0,0) and (2,3)

left to the top right. These images show how sensitive the basins of attraction are to the initial simplex used in Nelder-Mead. The basins of attraction are especially sensitive to the two assigned points of the simplex which will be important to note when studying different functions later in the paper.

Also, of all three images, we can note that the method does a relatively good job of finding a minimum at each point. There are only seven green colored dots within each image. This means the Nelder-Mead method failed to find a minimum only approximately 2% of the time in the grid that we were studying. So while the basin of attraction doesn't follow the pattern we would expect it to, it does narrow down on a minimum in almost every location of the grid.

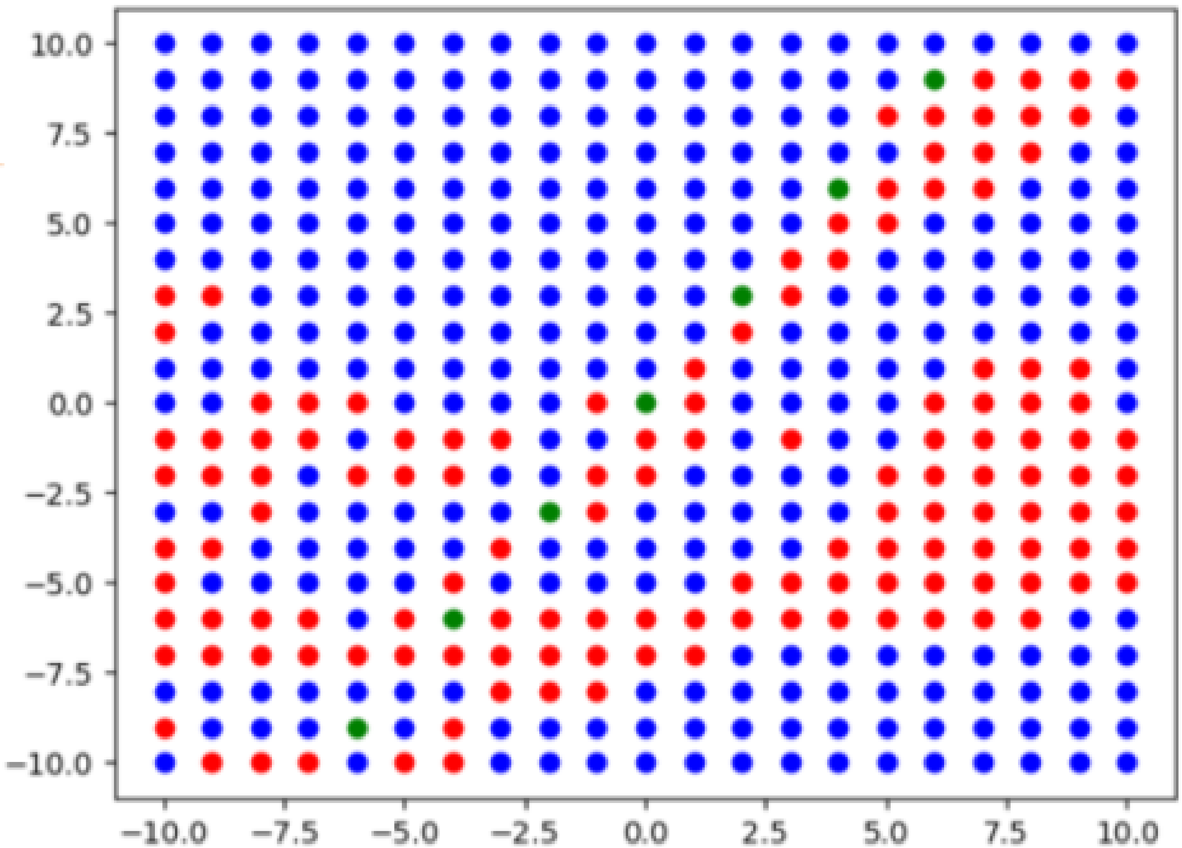


Figure 6.5: Systematic Point Method for $f(x, y) = 2x^2 - 4xy + y^4 + 2$ with assigned simplex points (0,0) and (-2,-3)

6.3 Randomized Centroid Method

After creating several different basins of attraction using the systematic point method, I created a basin of attraction using the randomized centroid method. As mentioned previously, this method created an initial simplex using three random points and then colored the centroid of that initial simplex based off of the minimum found. Once again, blue was used to symbolize (-1,-1) being found and red was used to symbolize (1,1) being found.

Once again, several different basins of attraction can be created using this method as each creation will result in different random points being selected to create the initial simplex. Figure 6.6 shows one iteration of the basin of attraction created using this method. In this image, I plotted 1000 different triangles to create the basin of attraction with the boundaries on the random points being between -10 and 10.

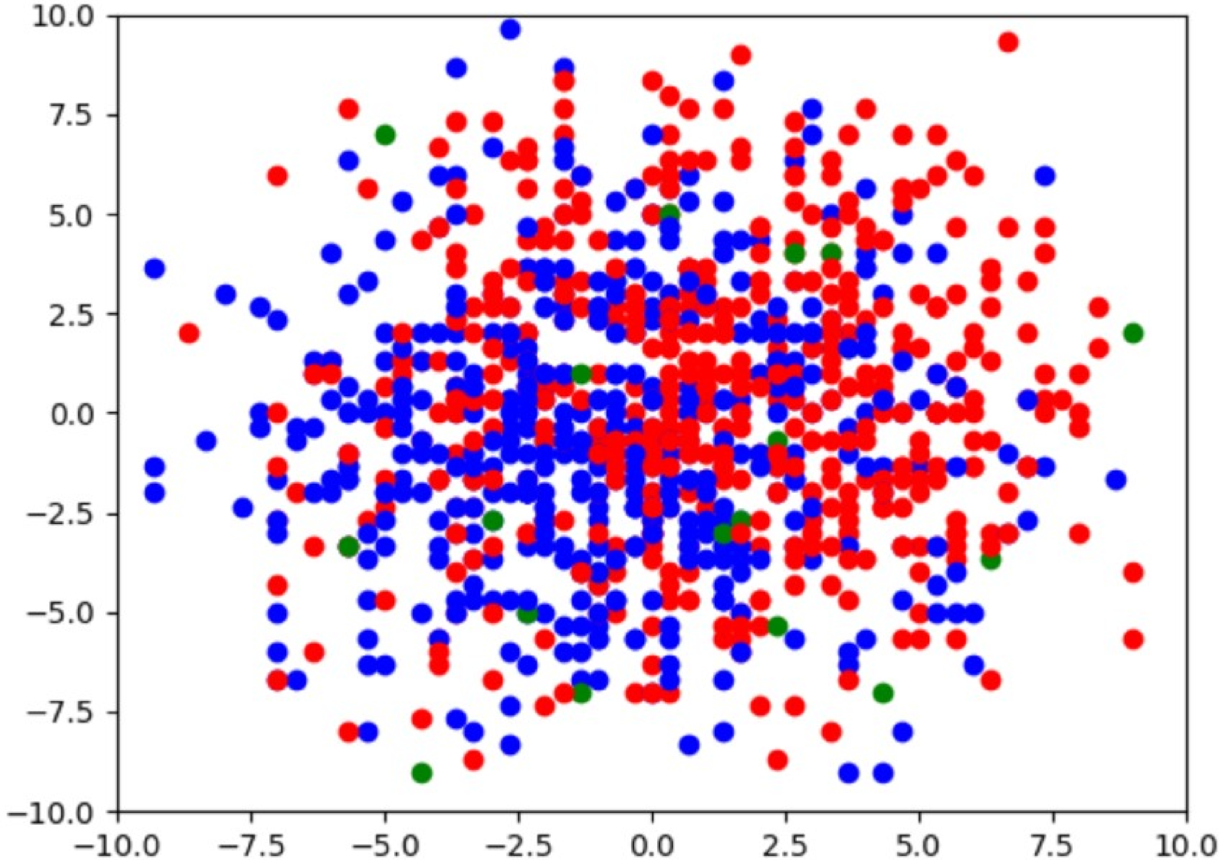


Figure 6.6: Random Centroid Method for $f(x, y) = 2x^2 - 4xy + y^4 + 2$

It's easy to note that it is extremely difficult to see what is happening in Figure 6.6. There does seem to be the appearance that we initially thought we would see where the top right corner has more red dots, signifying it found the minimum at (1,1) more frequently, and the bottom left corner has more blue dots, signifying it found the minimum at (-1,-1) more frequently. The red and blue dots seem to have an equal presence on the image which we confirm through their percentages. The minimum (-1,-1) was found approximately 47.2% of the time whereas the minimum (1,1) was found approximately 50.7% of the time. A minimum was not found 2.1% of the time.

While we mostly see the red and blue dots in their respective corners, we do see a scattering of the other colored dots as well. However, since we only used random points to create the simplices, we are unable to figure out why this is the case. Similarly, it is almost impossible to figure out why the green dots appear where they do and why the method failed to find a minimum when that particular point was the centroid of the initial simplex.

Overall, this method gave us the image that we were expecting to see but also failed to let us understand any of the results that we received. For this reason, we will not be using this method while studying future equations but will instead use the systemized centroid method, which will be discussed next, as it gives a much clearer picture on the basin of attraction for a function.

6.4 Systemized Centroid Method

The systemized centroid method once again implements the grid system which helps to make the basin of attraction images clearer and easier to read than the randomized centroid method. As mentioned previously, this method creates a simplex systematically picking points around a given point in the grid. I created simplices of various sizes by increasing or decreasing the values added around the given point. The real centroid is then calculated and colored in the final basin of attraction image.

Figure 6.7 shows a basin of attraction image using the systemized centroid method that shifts around the given point by one unit in each direction. This method, similar to the random centroid, shows us the image we expected to see with a bit more clarity. Most of the points on the bottom left found the minimum at (-1,-1) and most of the points on the top right found the minimum at (1,1). The minimum at (-1,-1) was found 53.29% of the time and the minimum at (1,1) was found 46.71% of the time. There were also no points at which the minimum was not found.

Just like the other methods, we notice that there are a scattering of red dots in the bottom left and a scattering of blue dots in the top right. This is interesting considering how small our initial simplex is. In a case like this, we would expect for there to be a complete diagonal split in the color of the dots. The best explanation for this is simply how the Nelder-Mead method "crawls" around the function to find it's minima. With these simplices, the method gets "caught" while moving toward the expected minimum and finds a path towards the other minimum instead.

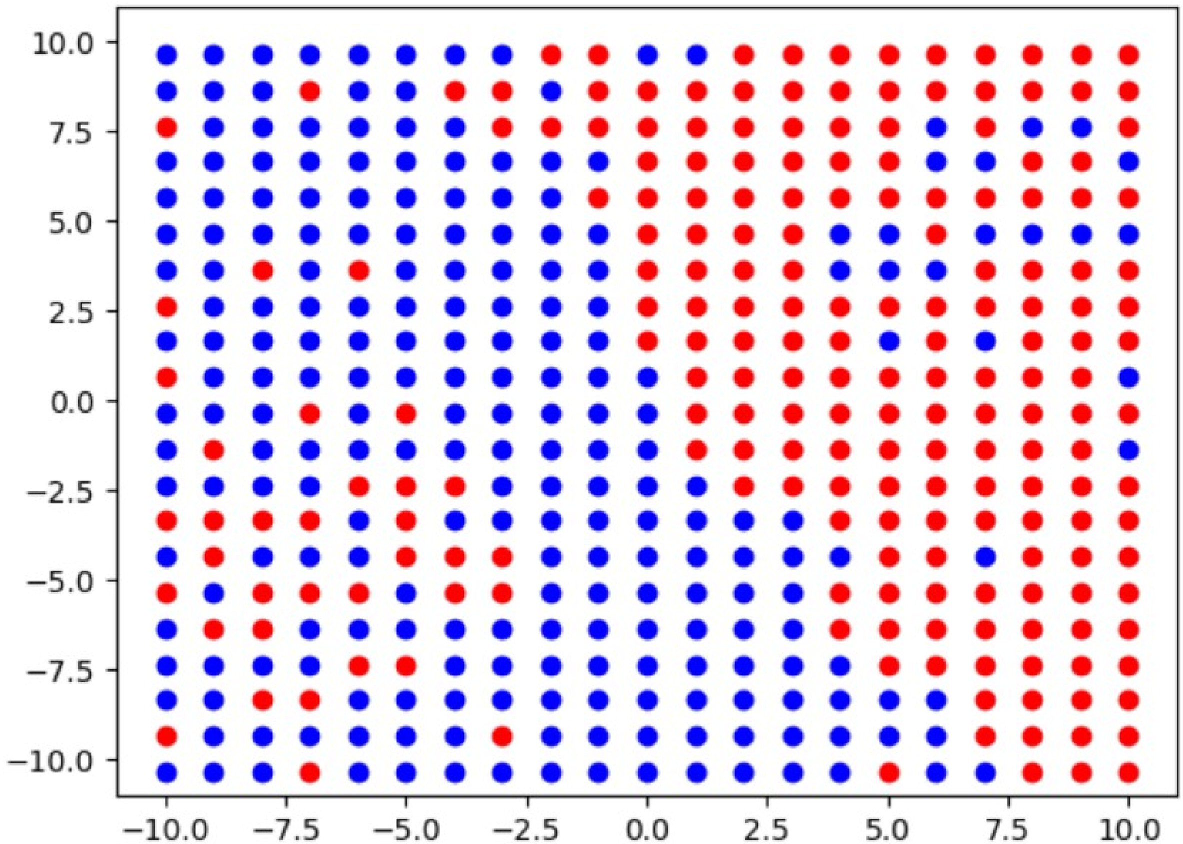


Figure 6.7: Systematized Centroid Method for $f(x, y) = 2x^2 - 4xy + y^4 + 2$ Shifting by One Unit

Figure 6.8 shows a basin of attraction for the same method but instead shifts around the given point from the grid by five units. Similarly, Figure 6.9 shows a basin of attraction with the simplices created by shifting around the given point from the grid by ten points. Notice that the axes have changed for these figures as we must recalculate the centroids for each simplex since they are now larger.

In both of these images, we see that the structure of the basin of attraction we saw in

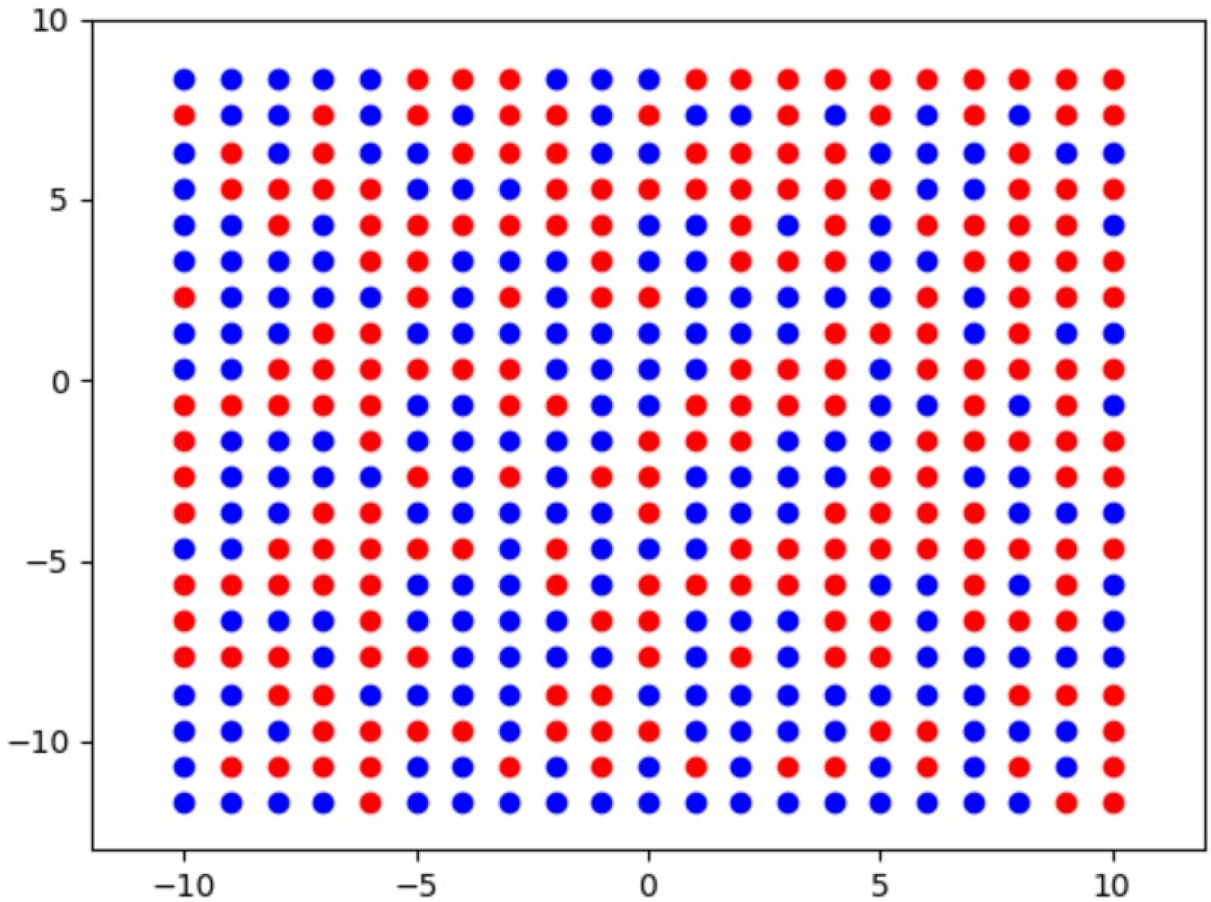


Figure 6.8: Systematized Centroid Method for $f(x, y) = 2x^2 - 4xy + y^4 + 2$ Shifting by Five Units

Figure 6.7 is gone. The coloring of points across the basin of attraction is now much more random. This is extremely interesting as we see that as our initial simplex gets larger, it becomes harder to predict which minimum a given simplex will find.

It is interest to note, however, that the larger simplices do not favor either minimum more than the other. When we shifted by five units, the minimum at (-1,-1) was found 50.34% of the time and the minimum at (1.1) was found 49.66% of the time. Similarly, when shifting by ten units, the minimum at (-1,-1) was found 51.02% of the time and the minimum at (1.1) was found 48.98% of the time. Once again, neither instance caused the method to stall or fail leading to zero green points on the basin of attraction image.

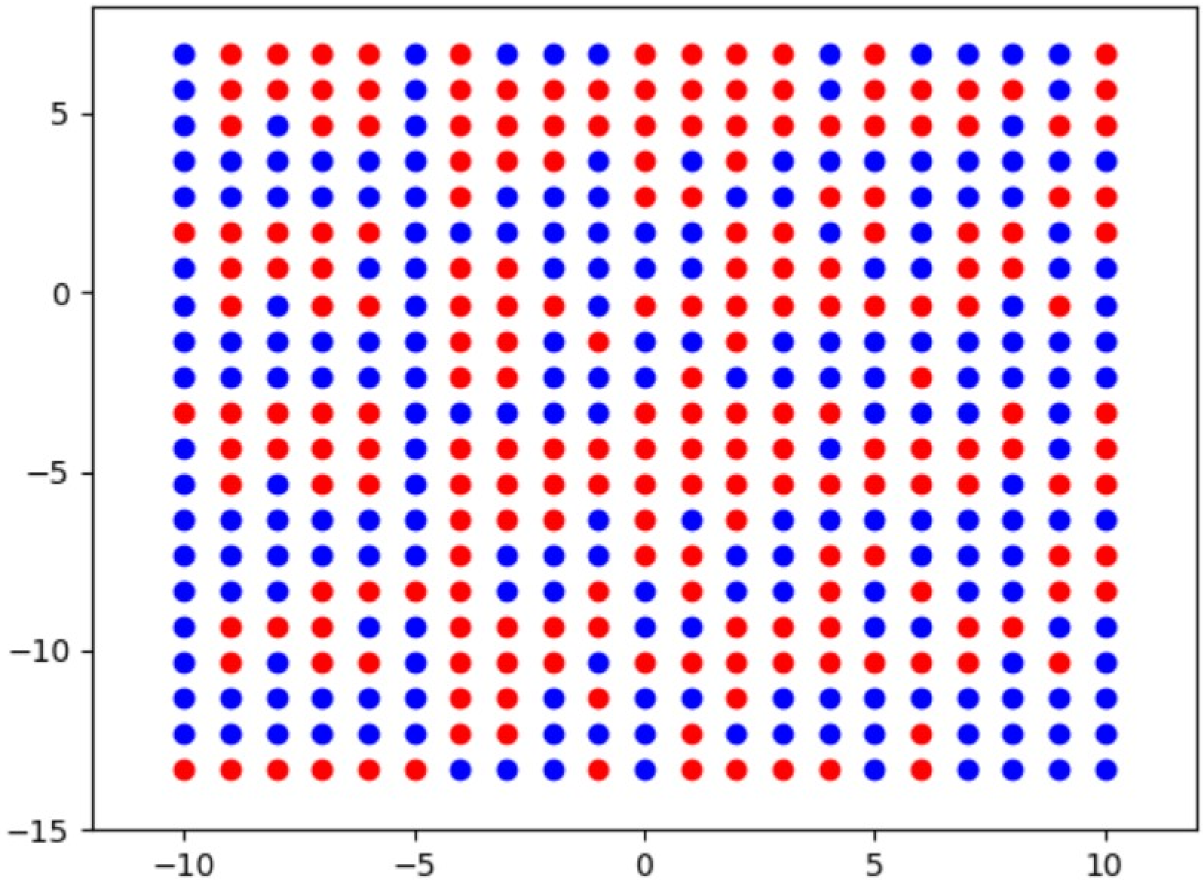


Figure 6.9: Systematized Centroid Method for $f(x, y) = 2x^2 - 4xy + y^4 + 2$ Shifting by Ten Units

6.5 Conclusions

Overall, we see that there are three methods that can be used to create basin of attraction images. The systematic point method finds minima with a relatively high accuracy. However, the minimum found in majority is extremely sensitive to the initial simplex that is used. In comparison, the randomized centroid method and the systemized centroid method also find minima with a high accuracy. Unlike the systematic point method, both methods manage to find both minima with an equal frequency.

Looking further into each method, however, we recognize that the randomized centroid method does not create basin of attraction methods that are easy to understand. We cannot tell which initial simplex creates the centroid that is color-coded and therefore cannot have a deeper understanding of the image. Meanwhile, we can gain a deeper understanding of the basin of attraction images from the systematic point method and systemized centroid methods. Using these two methods we can understand which simplices find particular minima. This helps us to understand the function that we are studying better as well.

Overall, the systemized centroid method seems to have the most accuracy at finding minima at the highest frequency and with an equal frequency for both minima. We noticed that as the simplices became larger using this method, the structures of the basin of attraction images were more loosely related to the contour plots of the function. However, despite the simplex size, neither minimum was favored more heavily than the other which cannot be said of the systematic point method. Therefore, the systemized centroid method seems to be the most stable in consistently and accurately finding minima using the Nelder-Mead method.

In the next chapters, we test this against different functions to see if it continues to hold true.

Chapter 7

Function with Global and Local Minimum

7.1 Understanding the Function

The next function studied had both a global and local minimum to understand how well the Nelder-Mead method could continue to find global minima when local minima were also present. This was the function studied:

$$f(x, y) = (x^2 + (y-1)^2)(x^2 + (y+1)^2) + \frac{1}{4}(x^2 + (y+1)^2)$$

There is only one global and one local minimum. The global minimum of this function is located at (0,-1) and the local minimum is located at (0,0.85). The graph, Figure 7.1, of the function helps to visualize it.

The following contour plots, Figure 7.2A and 7.2B also helps us to understand the function further. We can see that 7.2A shows the contour plot using the same axes we use to create the basin of attraction images. From this image, it is difficult to note the two minima. However, the zoomed in contour plot showed in Figure 7.2B helps us to see the global minimum at (0,-1) and the local minimum at (0,0.85).

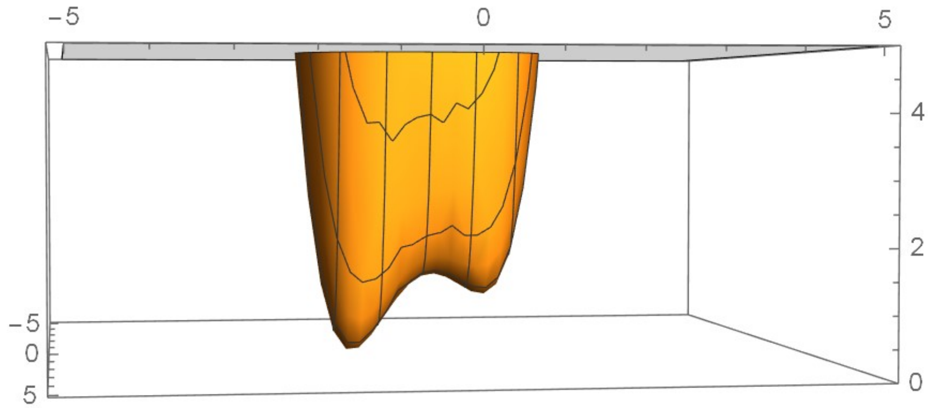


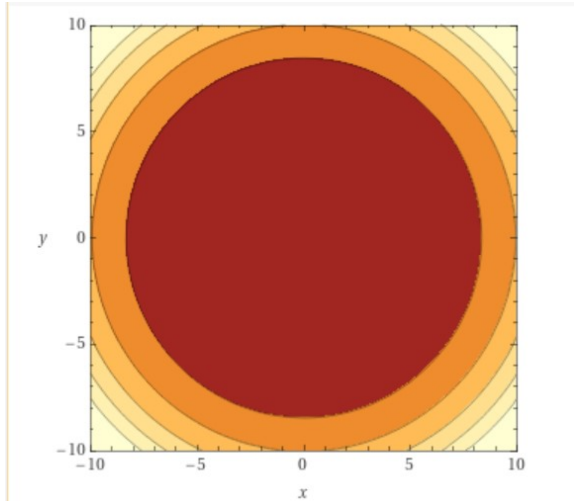
Figure 7.1: 3D Plot of $f(x,y) = (x^2 + (y-1)^2)(x^2 + (y+1)^2) + \frac{1}{4}(x^2 + (y+1)^2)$

Since we have a local and global minimum, we hope that the Nelder-Mead method finds the "true minimum", meaning the global minimum. However, since the method will not travel in an upward direction, it seems very likely that if the method finds the local minimum, it will stop there. Therefore, we would expect the basin of attraction image to find the local minimum when the initial simplex is near that point. Otherwise, we would expect to have the simplex find the global minimum.

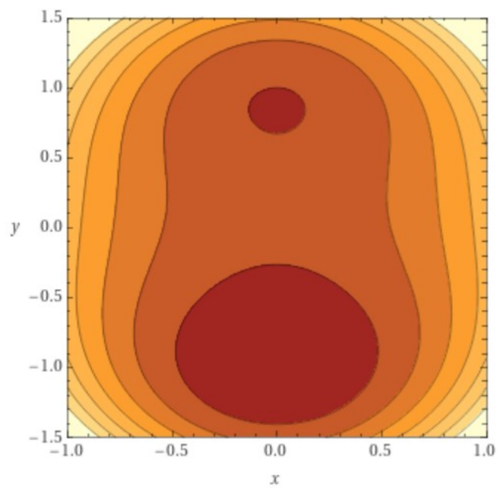
7.2 Systematic Point Method

We once again used the systematic point method to create the basin of attraction images. We use the same method to create the initial simplices but change the two assigned initial points used to build the simplex. In this case, we used initial points of (2,2) and (-2,-2). The third point was once again assigned using the previous method described. In this case, the color blue was used to denote the global minimum, (0,-1), and the color red was used to denote the local minimum, (0,0.85). Green is still used to denote that the initial simplex was unable to find a minimum.

Figure 7.3 shows images created using this method. We see that the graph has one again split itself into different quadrants. The top left and bottom right mostly found the local



(a) Zoomed Out Contour Plot



(b) Zoomed in Contour Plot

Figure 7.2: Contour Plots of $f(x, y)$

minimum whereas the bottom left and top right mostly found the global minimum. There is also a diagonal line from the bottom left to the top right of green points - meaning those points, when used as the systematic point in the initial simplex, were unable to find either minimum of the function. One reason for this may be that the initial points are also on that diagonal line and therefore the method may fail if the three simplex points are co-linear.

We see in this image, that we do not get the results we expected. The initial simplices in this basin of attraction image seem to have found the local minimum more often than the global minimum. The local minimum was found 63.04% of the time whereas the global

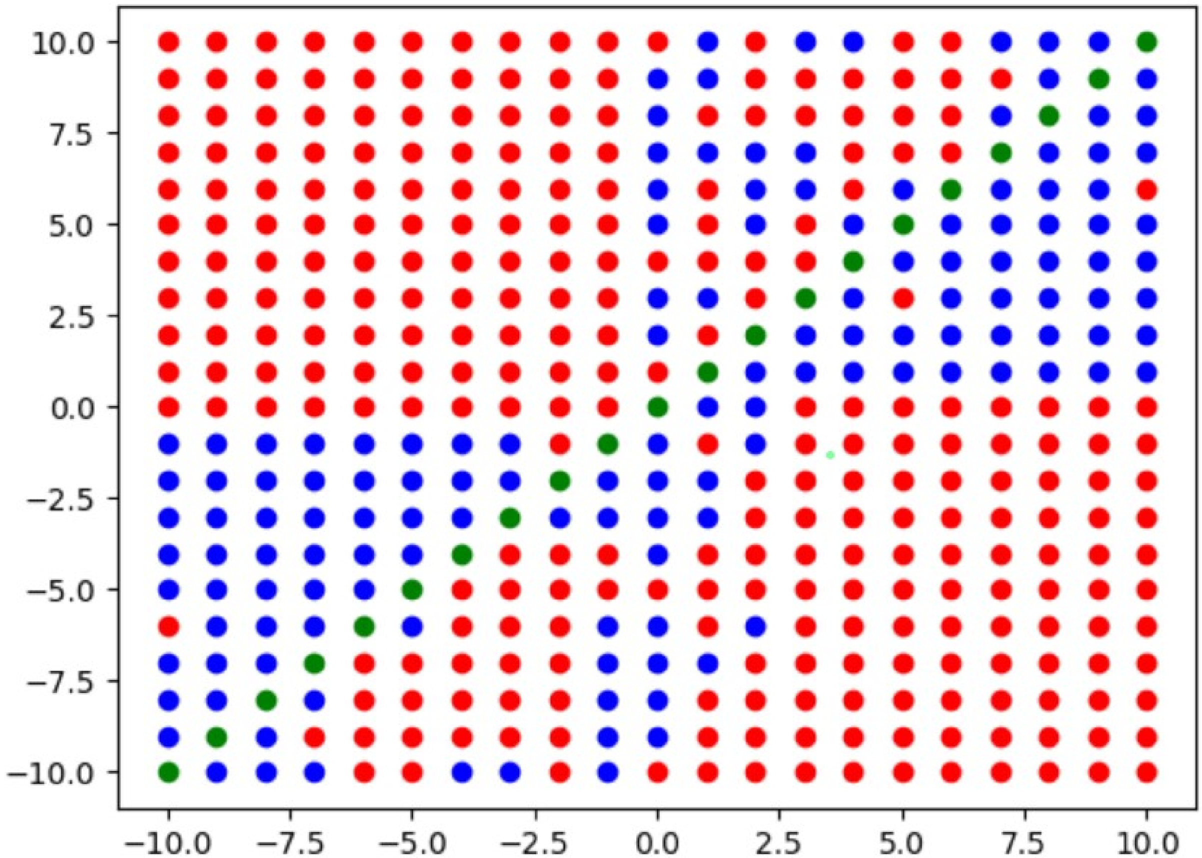


Figure 7.3: Systematic Point Method for $f(x,y) = (x^2+(y-1)^2)(x^2+(y+1)^2)+\frac{1}{4}(x^2+(y+1)^2)$ with assigned simplex points (2,2) and (-2,-2)

minimum was only found 32.2% of the time. An initial simplex did not find a minimum 4.76% of them. While using Nelder-Mead we would want the method to find the "true" minimum which is the global minimum. In this case, we do not see that happening. However, this is not the only basin of attraction we can create for this function.

When we change the initial assigned points to (0,0) and (-2,-2), there is a change in the percentage of points that find the global minimum and the local minimum. This is seen in image 7.4 where the global minimum was found 56.24% of the time and the local minimum was found 39% of the time. A minimum was not found 4.76% of the time once again.

Comparing these images further shows how the basin of attraction images using the

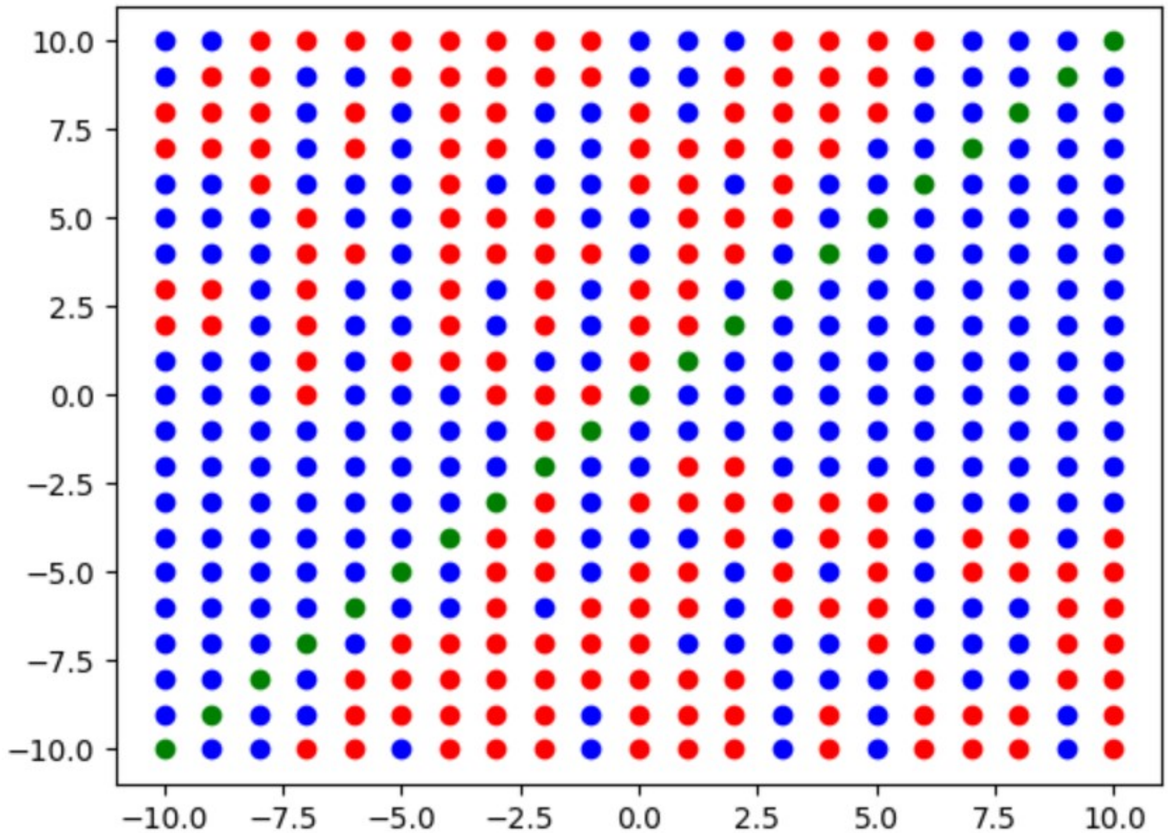


Figure 7.4: Systematic Point Method for $f(x, y) = (x^2+(y-1)^2)(x^2+(y+1)^2)+\frac{1}{4}(x^2+(y+1)^2)$ with assigned simplex points (0,0) and (-2,-2)

systematic point method are sensitive to the initial simplex used. We can see how the ability to find the global and local minimum changes after the assigned points used to create the initial simplex changes. If we want to mainly find the global minimum using the systematic point method, we have to use an initial simplex close to the global minimum. Similarly, initial simplices close to the local minimum will mainly "crawl" towards the local minimum. Without carefully selecting our initial simplex, we cannot guarantee finding a global minimum using the systematic point method and Nelder-Mead.

7.3 Systemized Centroid Method

As previously mentioned, we will not be studying the randomized centroid method as it is difficult to understand. Therefore, we study the systemized centroid method next. We once again look into how often the basin of attraction finds the global minimum compared to the local minimum using this method.

Figure 7.4 shows the image generated from using the systemized centroid method where we create the centroid by shifting around the grid point by one unit. The first thing that we note is that a minimum was always found in this basin of attraction as denoted by the lack of green points. However, it also seems that in this case, there doesn't seem to be any pattern in how the local or global minimum was found. There are no quadrants or sections artificially created by the red and blue points. However, we do note that the there do seem to be more blue points than red points in this basin of attraction. This means that the global minimum was found more often than the local minimum. In fact, the global minimum was found 80.27% of the time whereas the local minimum was found 19.73% of the time.

When we generate images for the systemized centroid by shifting around the grid point by five and ten units, we see a similar pattern. The basins of attraction created by shifting by five and ten units is depicted in Figures 7.4 and 7.5, respectively. In Figure 7.4, the global minimum was found 83.9% of the time and the local minimum was found 16.1% of the time. Meanwhile in Figure 7.5, the global minimum was found 81.86% of the time and the local minimum was found 18.14% of the time.

When using the systematic point method, we had to carefully create an initial simplex to find the global minimum more often than not. However, that does not appear to be the case with the systemized centroid method. It is quite reliable at finding the global minimum no matter the size or location of the initial simplex. We notice that at almost every size of the initial simplex, the global minimum was found approximately with the same percentage. Therefore, in this case, it seems the systemized centroid method is more useful than the

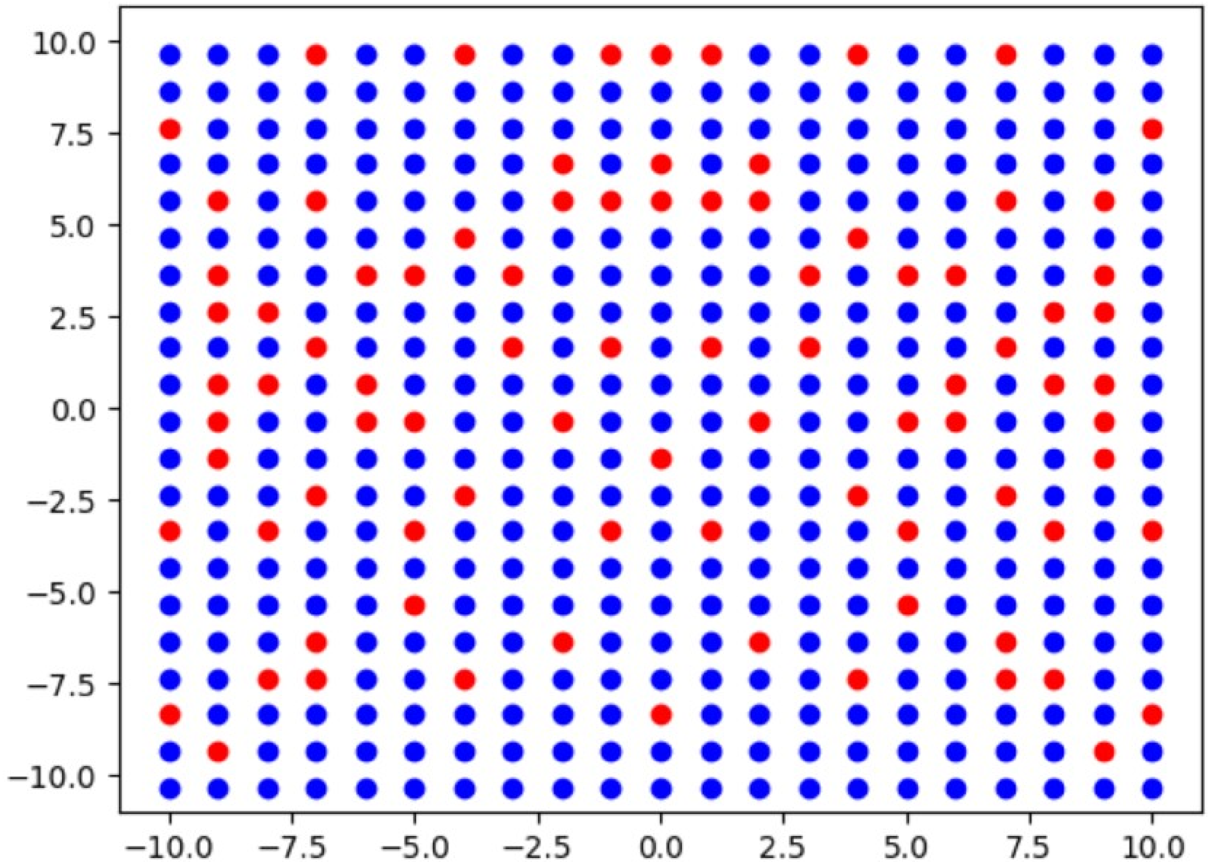


Figure 7.5: Systemized Centroid Method for $f(x, y) = (x^2 + (y-1)^2)(x^2 + (y+1)^2) + \frac{1}{4}(x^2 + (y+1)^2)$ Shifting by One Unit

systematic point method if we would like to find the global minimum.

7.4 Sensitivity Testing on the Function

In the previous section, we noted that the systemized centroid method was quite reliable at finding the global minimum for this function. However, it is also important to test if this would hold true if the function changed slightly. For example, if we kept the same global and local minima but the shape of the function changed, then would the basins of attraction also change?

We can use adjust the function we've been studying with coefficients in different ways as

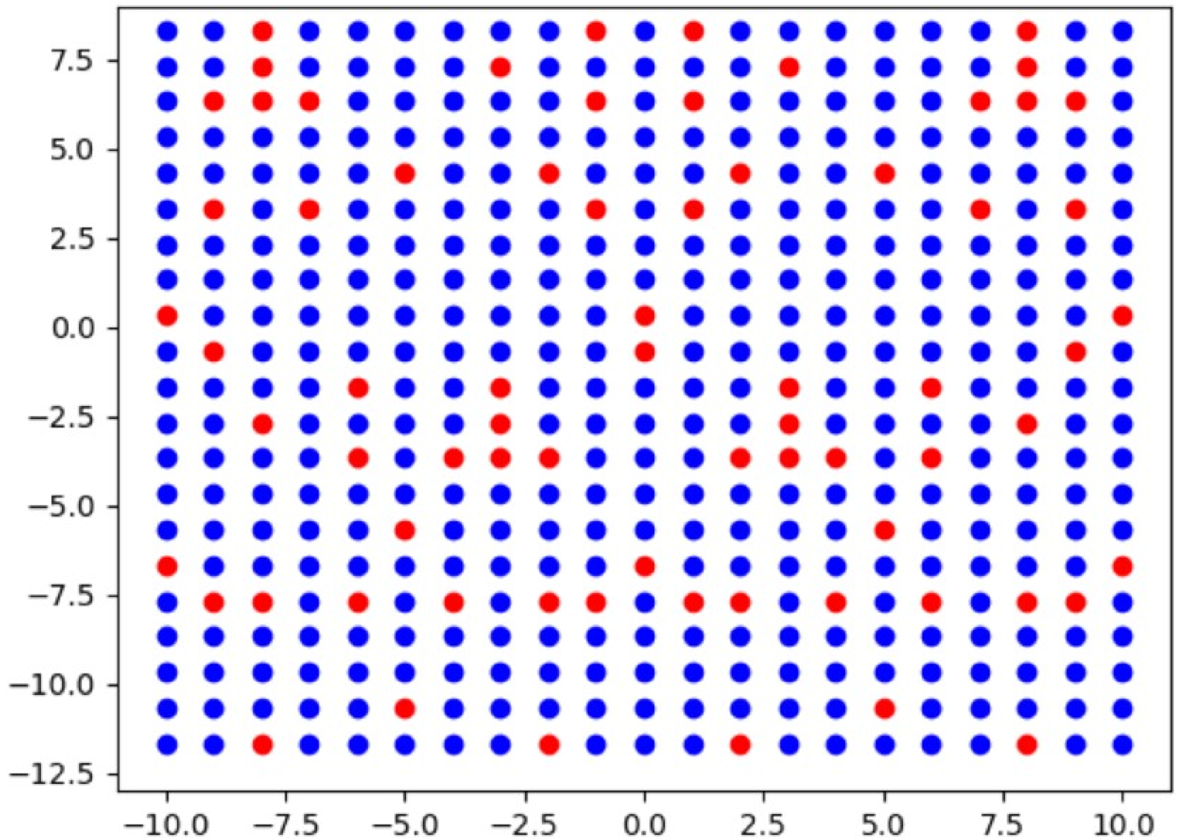


Figure 7.6: Systemized Centroid Method for $f(x, y) = (x^2 + (y-1)^2)(x^2 + (y+1)^2) + \frac{1}{4}(x^2 + (y+1)^2)$ Shifting by Five Units

seen below:

$$f(x, y) = (x^2 + (y-1)^2)(5x^2 + (y+1)^2) + \frac{1}{4}(5x^2 + (y+1)^2)$$

In this adjusted formulas, a coefficient of five has been added to certain terms that change the shape of the overall function but hold the same global and local minimum points. The graph and contour plot in Figure 7.8 and Figure 7.9, respectively, shows how the image has changed. While the graph may look similar to the original graph shown in Figure 7.1, we can see the differences in the original and new contour plots. The new function we are looking at is much narrower causing the function to have a steeper descent towards the minimum points. This means that in our basin of attraction images we expect to see the Nelder-Mead

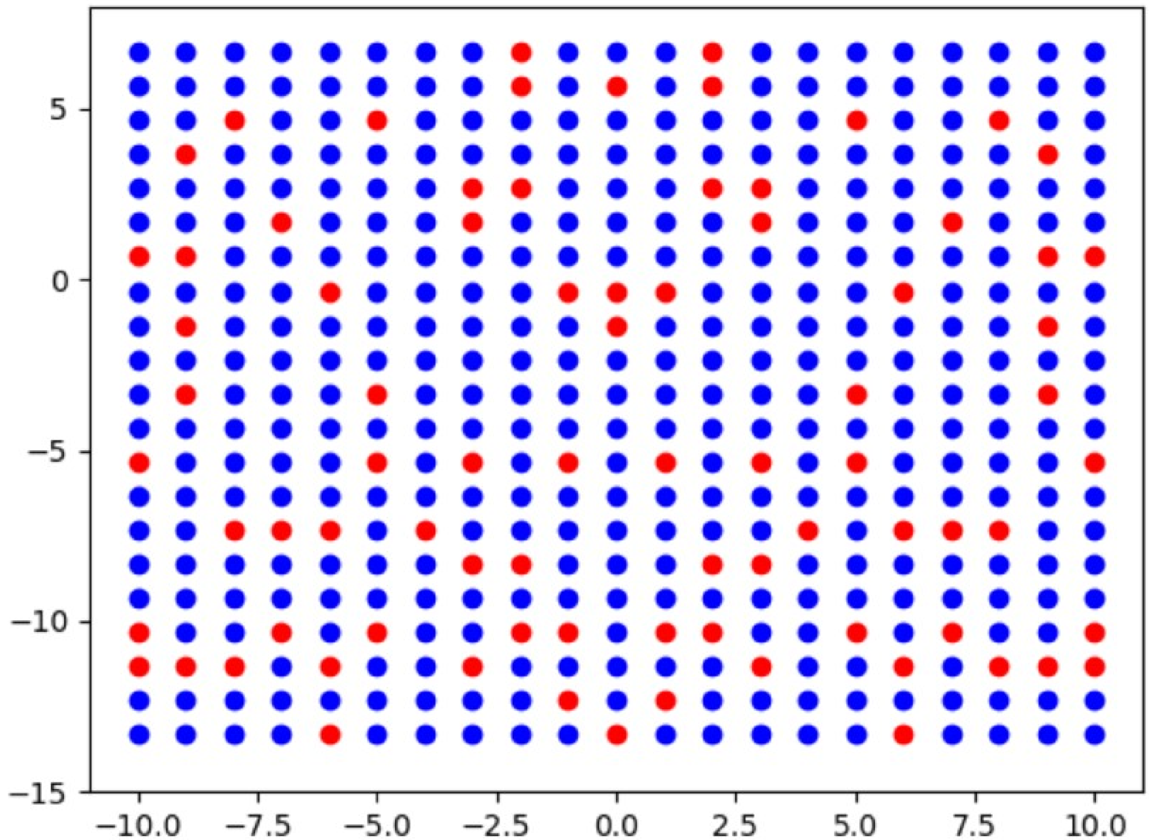


Figure 7.7: Systemized Centroid Method for $f(x,y) = (x^2 + (y-1)^2)(x^2 + (y+1)^2) + \frac{1}{4}(x^2 + (y+1)^2)$ Shifting by Ten Units

method find the local minimum more often than the global minimum as it will be easier to rapidly approach the local minimum at every step.

We can see this occurring in Figure 7.10 where we created a basin of attraction using the systemized centroid method while shifting by one unit on the adjusted function. There still does not appear to be a pattern between how the local and global minimum points are found, however, we do see that there are more red points in this basin of attraction image than we had previously seen in Figure 7.5. On the basin of attraction for the adjusted function, the global minimum was found 46.49% of the time and the local minimum was found 53.51% of the time. This means that the local minimum using the adjusted function was found 33.78% more than the local minimum in the original function.

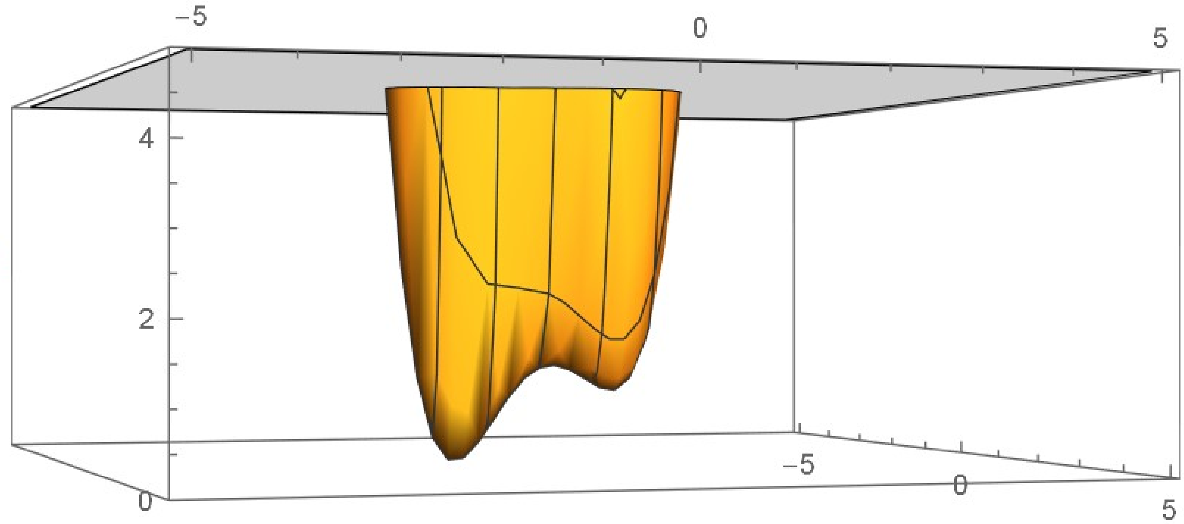


Figure 7.8: 3D Plot of $f(x,y) = (x^2 + (y-1)^2)(5x^2 + (y+1)^2) + \frac{1}{4}(5x^2 + (y+1)^2)$

We see this same phenomena occur when we create a basin of attraction for the adjusted function using the systemized centroid method while shifting by ten units. The global minimum is found 57.6% of the time and the local minimum is found 42.4% of the time. This means that the local minimum was found 24.26% more with the adjusted function than the original function. There is also an over 10% increase in the number of times the global minimum is found when we use a large simplex. This could mean that when the function has a steeper gradient of descent a larger simplex could help Nelder-Mead to find a global minimum rather than a local minimum.

Overall, it appears that the ability for Nelder-Mead to find a global minimum versus a local minimum is affected by the shape of the function. When we had a steeper function, the Nelder-Mead method found the local minimum more often than with a flatter function. It appears that the steeper descent forced the method to take larger steps in one direction thereby "trapping" it within the lower minimum rather than allowing it to crawl toward the global minimum when possible.

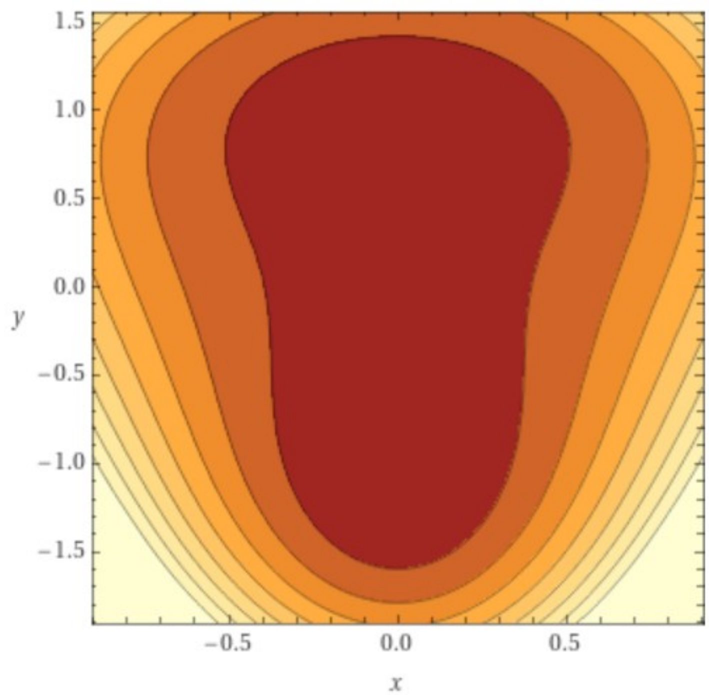


Figure 7.9: Contour Plot of $f(x,y) = (x^2 + (y-1)^2)(5x^2 + (y+1)^2) + \frac{1}{4}(5x^2 + (y+1)^2)$

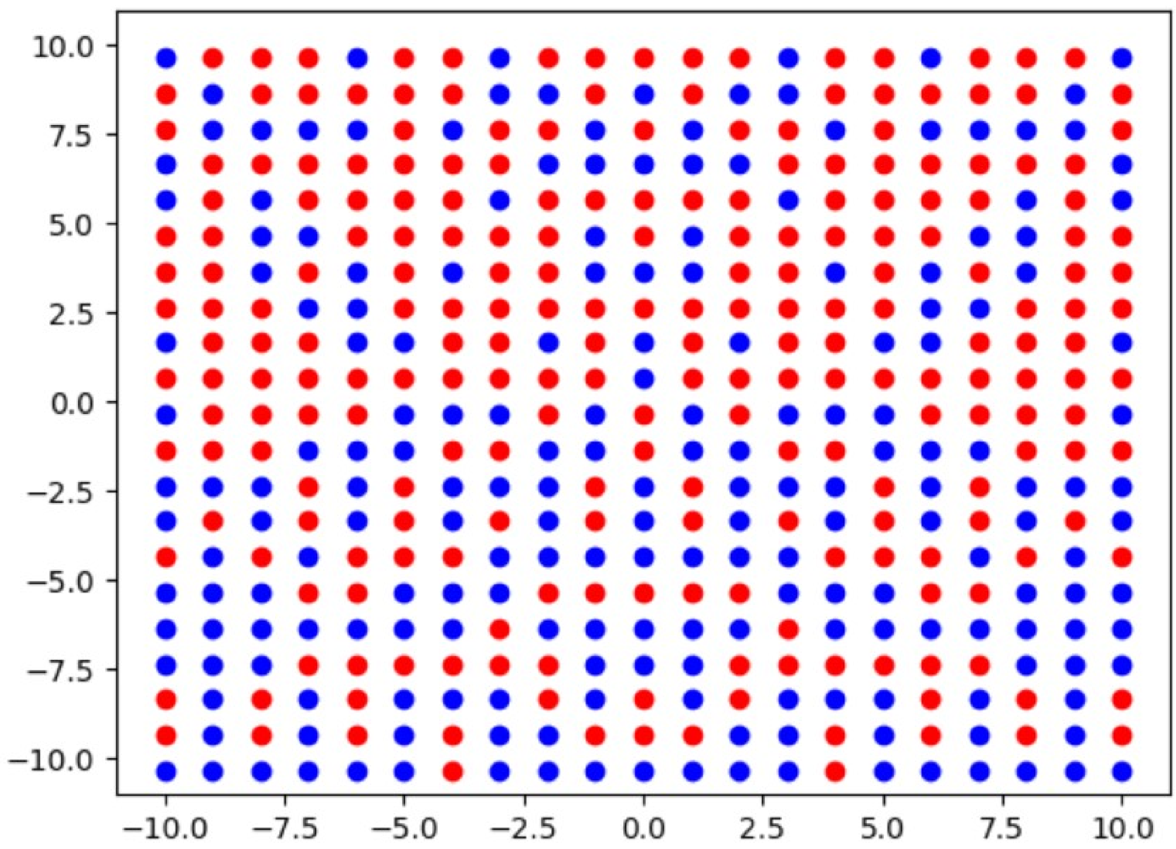


Figure 7.10: Systemized Centroid Method for $f(x,y) = (x^2 + (y-1)^2)(5x^2 + (y+1)^2) + \frac{1}{4}(5x^2 + (y+1)^2)$ Shifting by One Unit

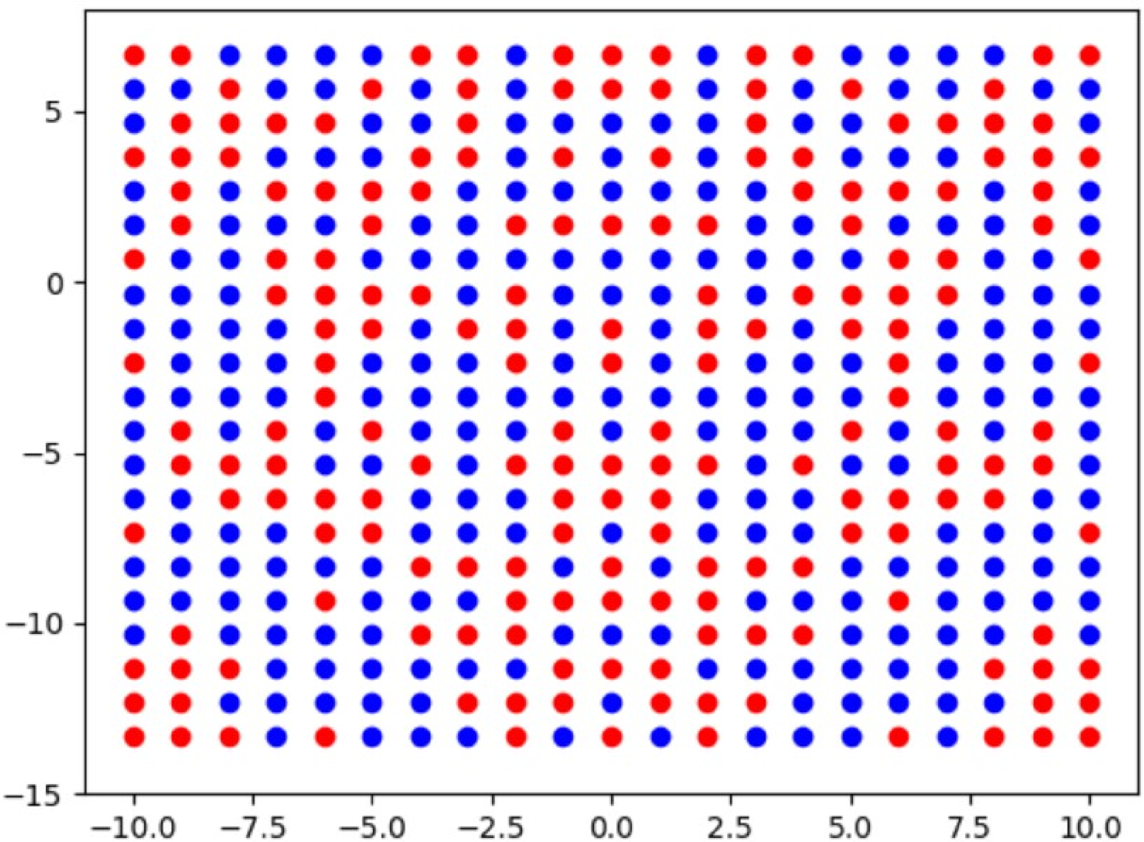


Figure 7.11: Systemized Centroid Method for $f(x, y) = (x^2 + (y-1)^2)(5x^2 + (y+1)^2) + \frac{1}{4}(5x^2 + (y+1)^2)$ Shifting by Ten Units

Chapter 8

Function with Multiple Minima

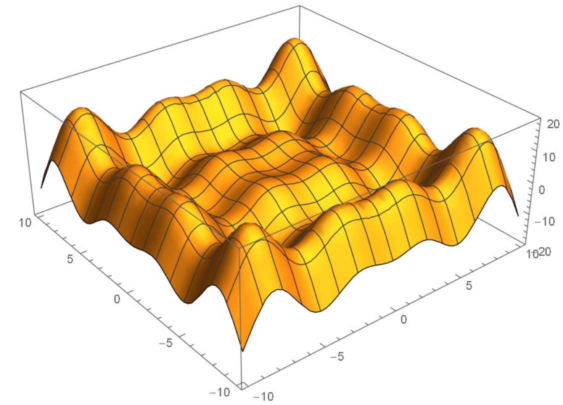
8.1 Understanding the Function

So far, we have noticed that Nelder-Mead is mostly very stable with finding minima. It can find any minimum a high percentage of the time and using the systemized centroid method is effective at finding a global minimum compared to a local minimum. Therefore, we next study a function that has multiple minima to ensure that the Nelder-Mead method remains stable - or continues to find minima, specifically global minima. The function studied is below:

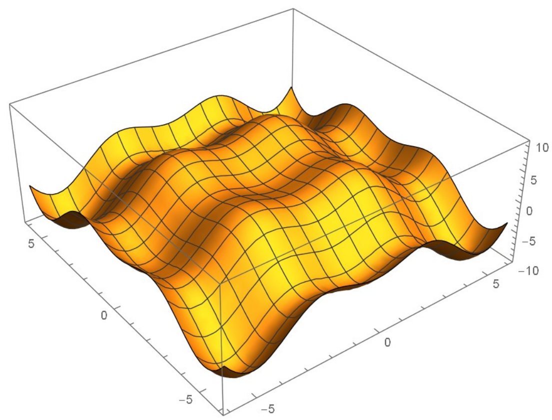
$$f(x, y) = xsin(x) + ysin(y)$$

We can see two different graphs of this function in Figure 8.1. Figure 8.1A shows the graph of the function on a plot with an x-range from [-10,10] and a y-range of [-10,10]. However, with these axes, the graph has certain maximums and minima near (0,0) but far away from that only moves in the downward direction. Therefore, we wouldn't have a global minimum that we could use to test Nelder-Mead. In order to be able to artificially create a global minimum, we create a smaller range as seen in Figure 8.1B.

In Figure 8.1B, we set an x and y domain of [-6,6]. Using this range, we have local



(a) Zoomed Out Graph



(b) Zoomed in Graph

Figure 8.1: Graphs of $f(x, y) = xsin(x) + ysin(y)$

minimum at (0,0) for which the function value is 0. We also have local minima at (0,4.9), (0,-4.9), (4.9,0), and (-4.9,0) for which the function value at these points is approximately -4.81. The global minima for this function are found at (4.9,4.9), (-4.9,4.9), (4.9,-4.9), and (-4.9,-4.9) with the function value at these points being approximately -9.63.

We can further understand this function by studying the contour plot shown in Figure 8.2. The function has many maximums and minima. We can see with the limited domain how the minima fall at the corners of the graph. In the middle of the edges and the center are the local minima of the graph.

Given the graph and the contour plot, we would hope that the a majority of the basin of attraction images find the global minima in the corner. As we saw in the last chapter,

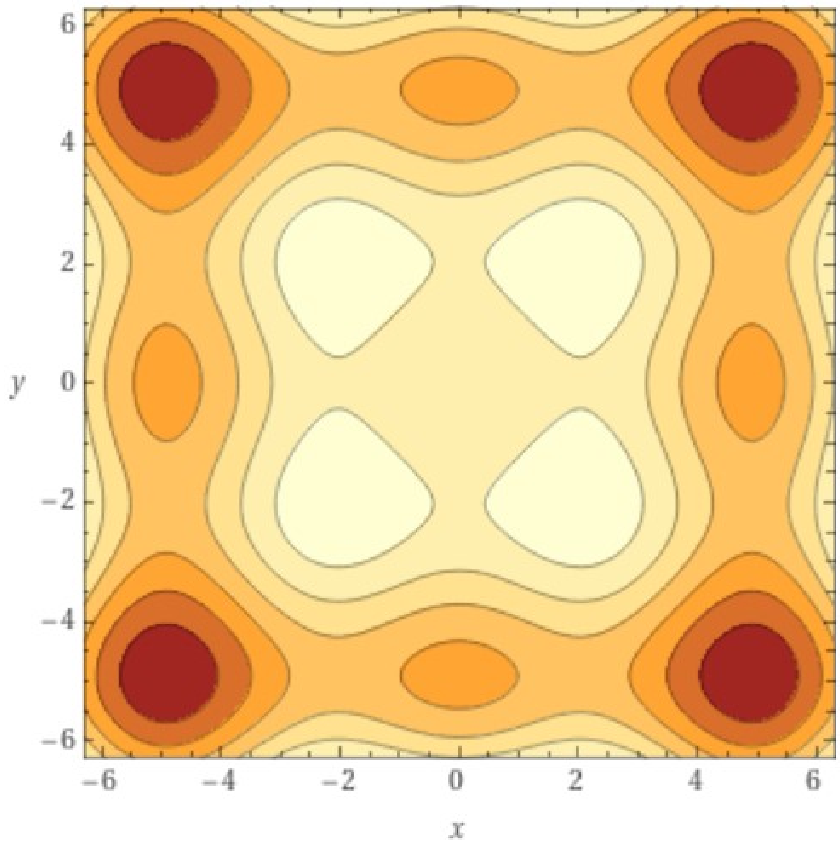


Figure 8.2: Contour Plot of $f(x, y) = xsin(x) + ysin(y)$

the global minimum has a steeper gradient making it seemingly easier for Nelder-Mead to be attracted to that minimum with the systemized centroid method. However, due to the large number of global and local minima within this function, it would not be surprising for the method to get "stuck" within a local minimum.

As we will see in the next few sections, each minimum will be given it's own color to identify which basin is found. For the global minima: (4.9,4.9) is colored purple, (-4.9,4.9) is colored orange, (4.9,-4.9) is colored yellow, and (-4.9,-4.9) is colored red. For the local minima: (0,0) is colored blue), (4.9,0) is colored pink, (-4.9,0) is colored gray, (0,4.9) is colored black, and (0,-4.9) is colored cyan. Green is still used to represent that a simplex was unable to find a global or local minimum.

8.2 Systematic Point Method

Once again, while using the systematic point method to create a basin of attraction, we must assign new initial starting points for the simplex. To start with, we use the points (1,1) and (-1,-1). The basin of attraction created by using these values to create the initial simplex is shown in Figure 8.3.

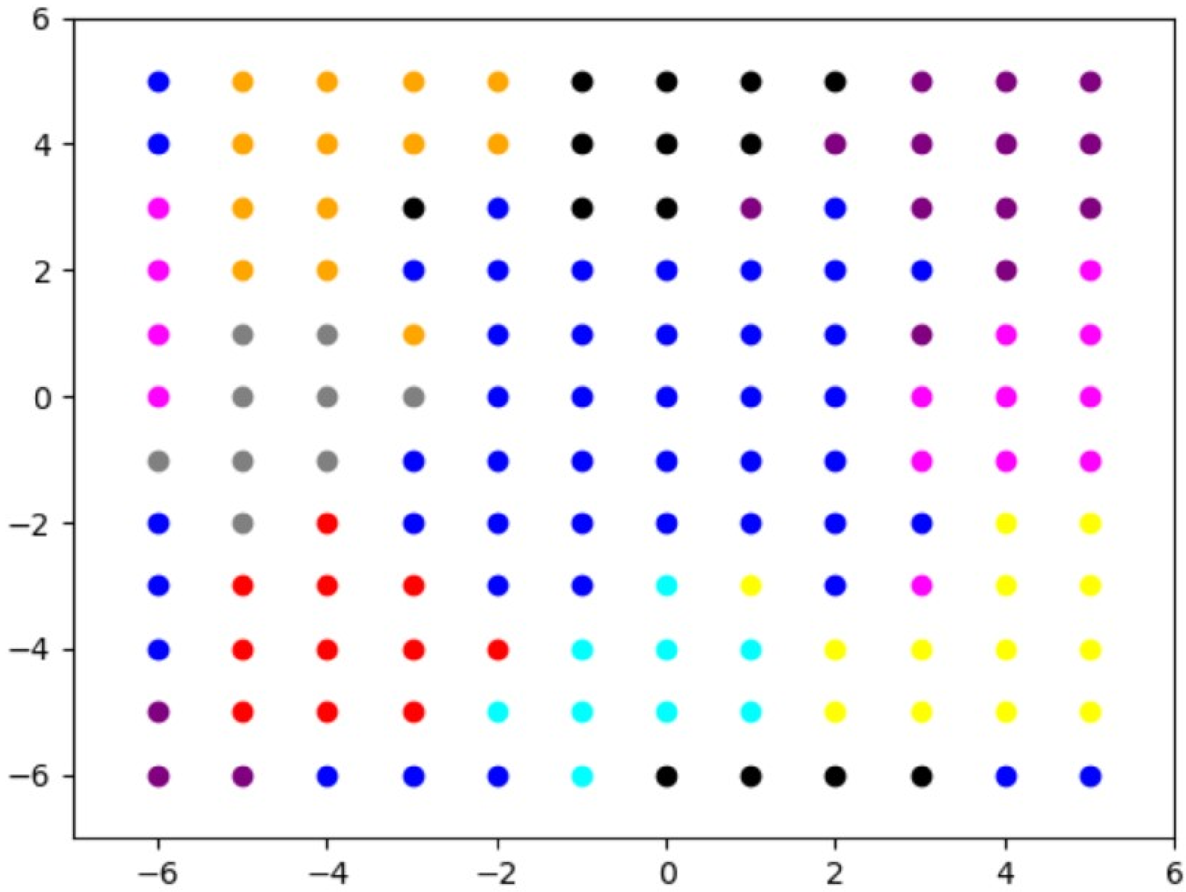


Figure 8.3: Systematic Point Method for $f(x, y) = xsin(x) + ysin(y)$ with assigned simplex points (1,1) and (-1,-1)

While using this particular initial simplex method, we can see that the majority of the time, the minima found is (0,0) - in fact, 31.25% of the time. This minimum was found mainly when the systematic point was located in the center of the image. Around the edges, we can see different colored dots symbolizing that other minima have been found. Most of the systemized points are colored the same as the minimum point they are closest to around the edges. The minimum at (4.9,4.9) is found 11.11% of the time which is the second highest percentage for a minimum being found. All other minima were found less than 10% of the time with the minimum at (0,4.9) being found the least at only 6.25% of the time.

We can also note the differences in the number of times global minima were found compared to local minima. Global minima were found 36.81% of the time. This means that local minima were found 63.19% of the time. Therefore, overall, using the systematic point method and these values for the initial simplex, we can see that the local minima are found almost twice as often as the global minima. We can change our initial simplex to see how that affects the number of global and local minima found.

In Figure 8.4, we have changed the assigned values of the simplex to be (3,3) and (-3,-3). In this case, we see the same pattern as we did in Figure 8.3. Most of the systemized points are colored according to the minimum they were found closest to. The minimum that was found the most was once again (0,0) and was found 29.86% of the time. The minimum found the least was still (0,4.9) along with (4.9,0) and both minima were found 6.25% of the time. Overall, the global minima were found 40.96% of the time and the local minima were found 59.02% of the time. Therefore, we see that this change in the initial simplex doesn't change how often the global minimum is found compared to the local minima. We can next check to see if a more drastic change to the initial simplex will find the global minima more frequently.

For Figure 8.5, we use assigned values of (-2,5) and (5,-2) to create the initial simplex. Using this simplex, we can see that the basin of attraction image has changed drastically from Figure 8.3 and 8.4. To begin with, we see that a minimum was not found from an initial simplex at six different points. We also see that rather than the points mainly being color coded based on the minimum they are closest too, there is a large uptick in the orange and black points which correspond to the minima at (-4.9,4.9) and (0,4.9), respectively. The

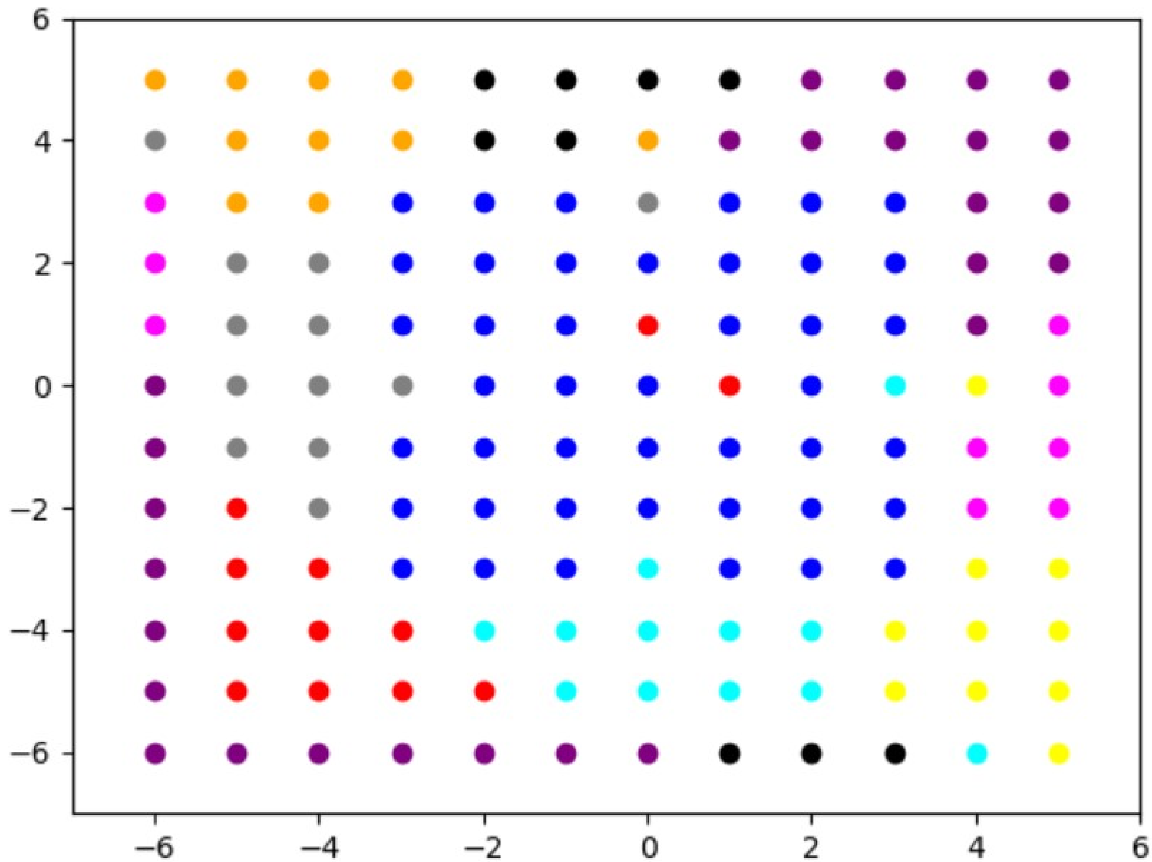


Figure 8.4: Systematic Point Method for $f(x,y) = xsin(x) + ysin(y)$ with assigned simplex points (3,3) and (-3,-3)

other minima are found more rarely in this basin of attraction image.

Overall, the minimum at (-4.9,4.9) was found the most at 39.58%. The minimum at (0,4.9) was found 29.86% of the time. In this image, the minimum at (0,0) was surprisingly never found which is different from the other basin of attraction images generated. Also, in this case, the global minima were found 54.87% of the time whereas the local minima were found 40.96% of the time. A minimum was not found 4.17% of the time. We can see that the global minima were found more often in this case than the previous basins of attraction images. Using this more drastic change in initial simplex, we are also able to find global minima more than local minima for the first time.

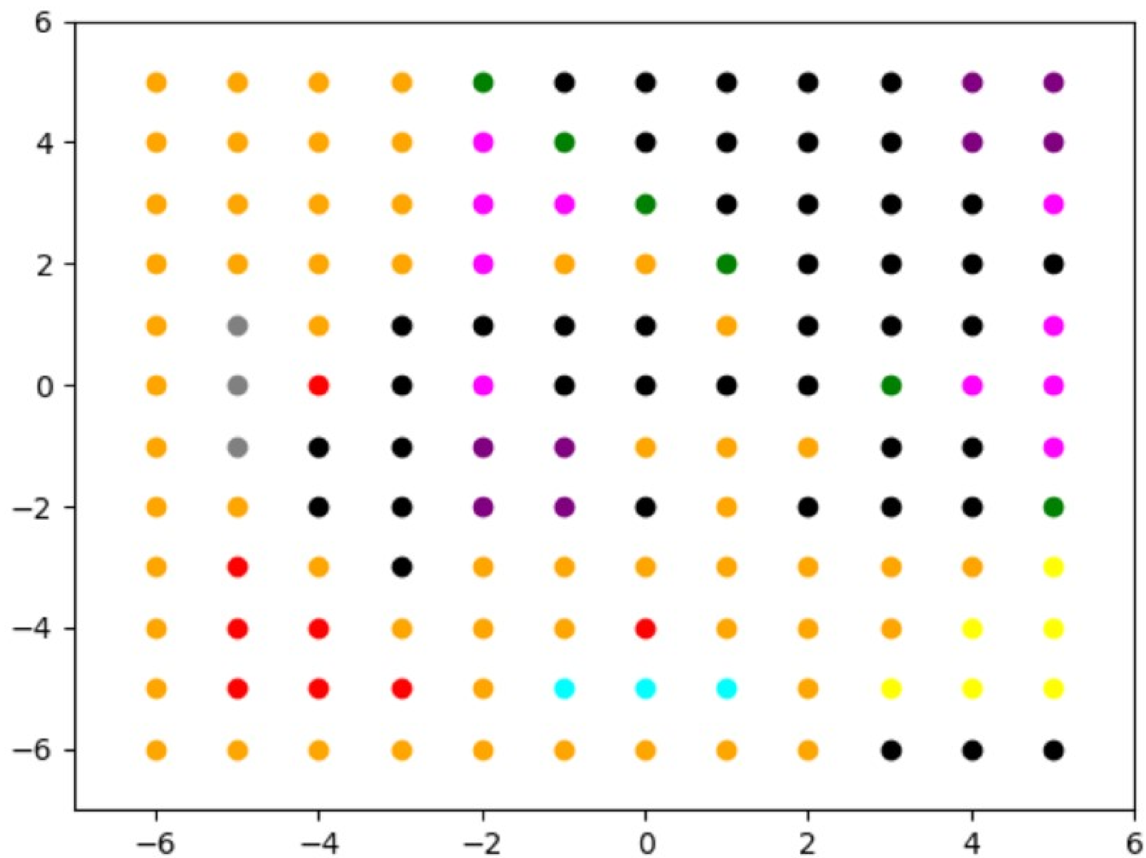


Figure 8.5: Systematic Point Method for $f(x, y) = xsin(x) + ysin(y)$ with assigned simplex points (-2,5) and (5,-2)

8.3 Systemized Centroid Method

As we saw in the last chapter, the systemized centroid method seemed to be more effective at finding global minima in certain cases. We attempt to see if it can be more effective on this function as well.

We start by creating the basin of attraction where we create the centroid by shifting around the grid by one unit as seen in Figure 8.6. As we can see in the image, the color of the points in this case are even more coordinated with the location of the minima they are closest to. However, unlike the randomized point method, we do not see a majority of any colors or minima found. The minimum found most was (-4.9,-4.9) and was found 17.36% of the time and the minimum found least was (0,4.9) and was found 6.25% of the time. We can see that this is a much tighter range in finding each minimum than there appeared to be within the systematic point method.

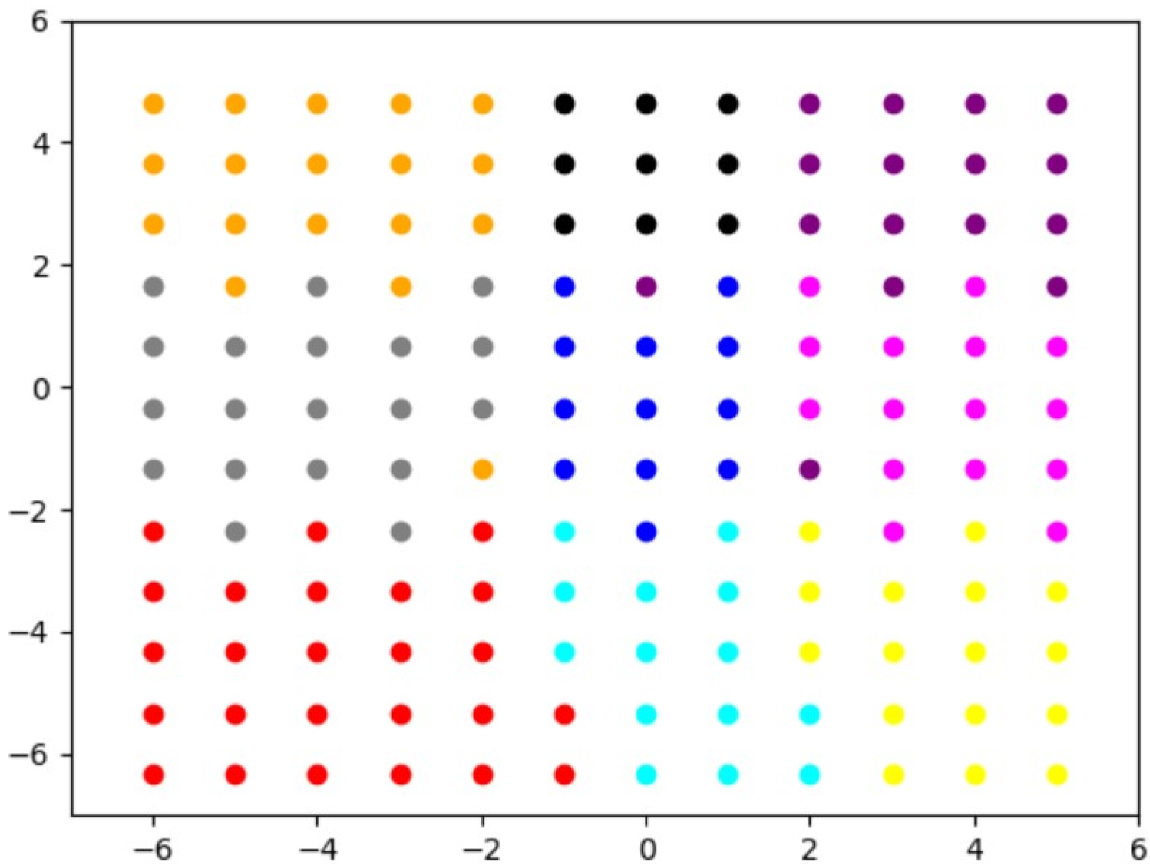


Figure 8.6: Systemized Centroid Method for $f(x, y) = xsin(x) + ysin(y)$ Shifting by One Unit

We can also compare the number of times global minima were found compared to local minima using this method. Global minima were found 52.08% of the time meaning that local minima were found 47.91% of the time. Therefore, the global minima were found more often than the local minima when using a tight range. Note that the only slightly more global minima were found than local minima which is not ideal. We also see that the four global minima in this case were found with an approximately equal frequency. In Figure 8.5, when we used the systematic point method and found the global and local minimum at approximately the same rate, the rates were mainly due to certain global and local minima rather than all of them. Therefore, we are finding the global minima we would expect to find with more precision.

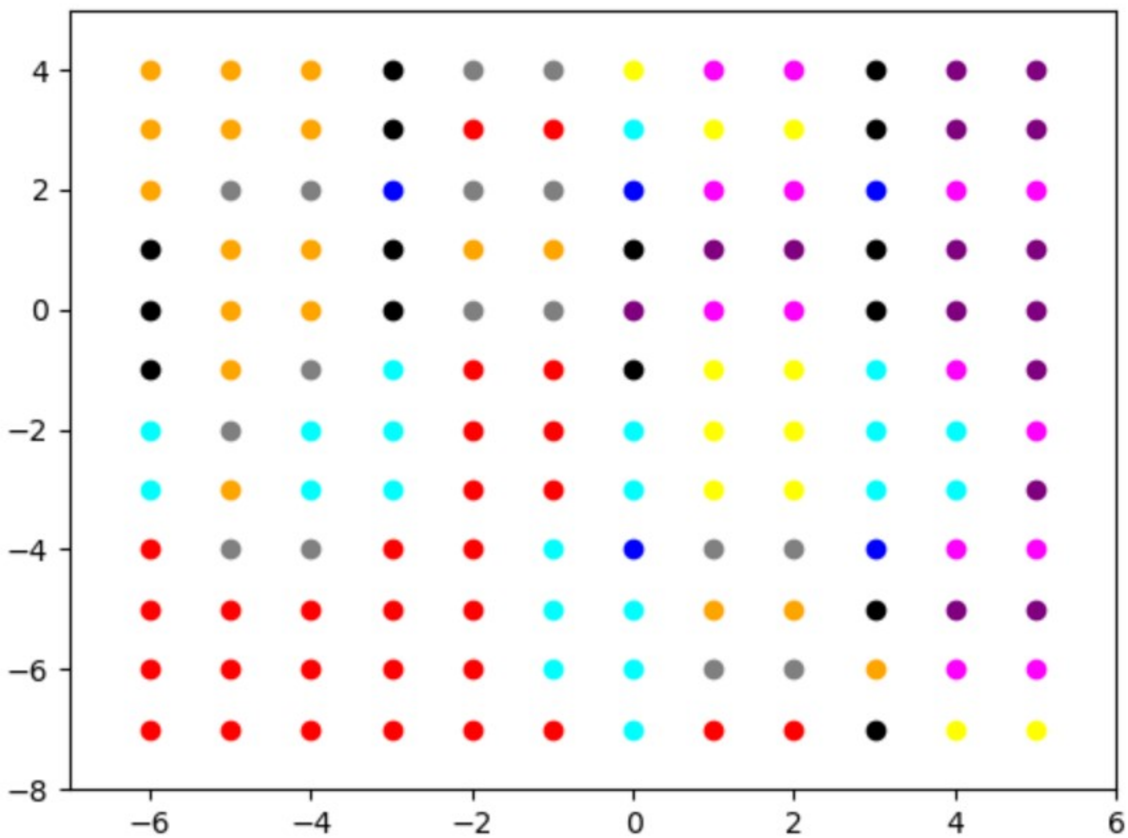


Figure 8.7: Systemized Centroid Method for $f(x, y) = xsin(x) + ysin(y)$ Shifting by Three Units

In Figure 8.7, we create the basin of attraction by shifting around the grid by three units to find the centroid. In this case, we can see that the preciseness of the minima being color coordinated according to location that we saw in Figure 8.6 is no longer present. Instead, the color coordinated points are very loosely around the minimum we would expect them to find. However, similar to Figure 8.6, in the basin of attraction depicted in Figure 8.7 the minimum (-4.9,-4.9) was found the most at 20.14%. The minimum at (0,0) was found the least, only 3.47% of the time. This makes sense as (0,0) is the shallowest minimum of this function and therefore that makes it easier for Nelder-Mead to not locate it as a minimum.

In this basin of attraction image, the global minima are found 50.7% of the time and the local minima are found 49.3% of the time. We can see that when the simplex is created with a larger shift, and is therefore larger itself, finding global and local minima becomes more comparable. Both the global and local minima are found with an almost identical frequency. We can test this out further by once again expanding the shift in points used to create the initial simplex.

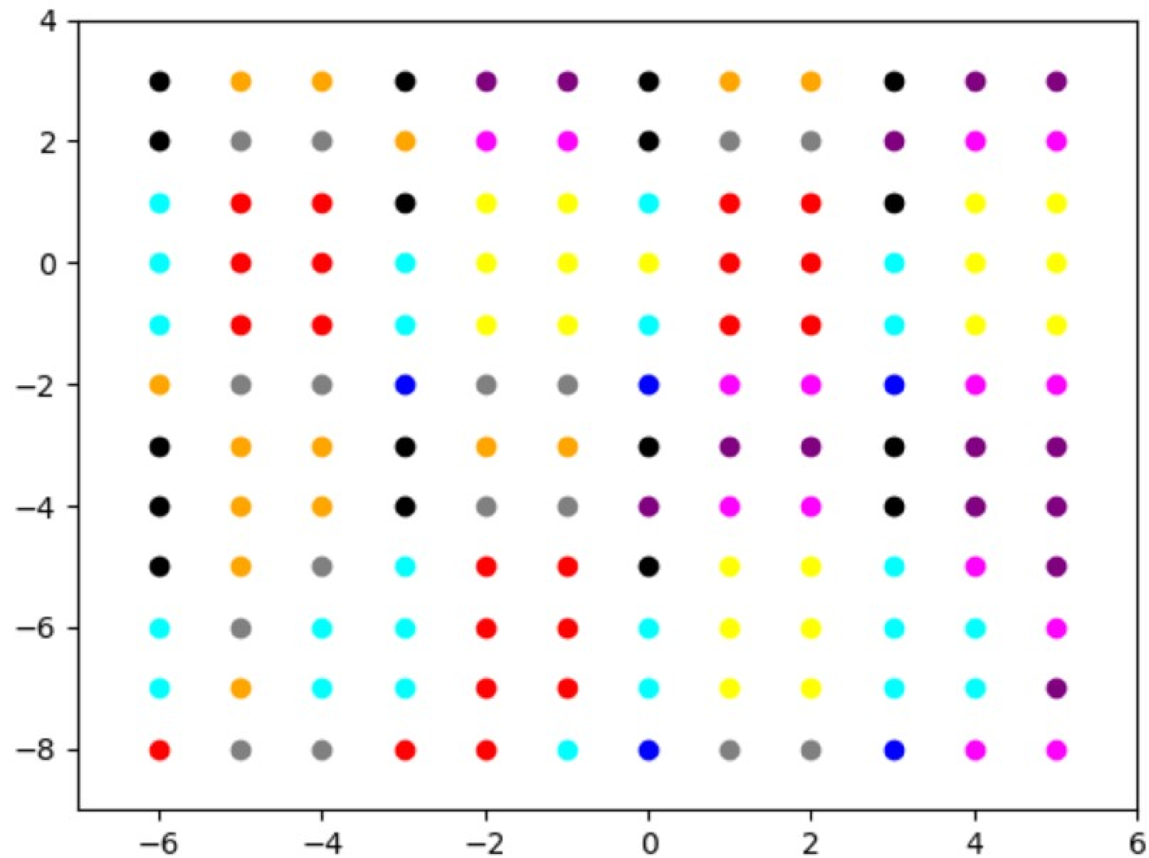


Figure 8.8: Systemized Centroid Method for $f(x, y) = xsin(x) + ysin(y)$ Shifting by Six Units

In Figure 8.8, we find the centroid of the initial simplex by shifting around the grid point by 6 units. In this case, we can see that the color coordination of the points falls even more loosely around the location we would expect them to - or than where they were located in

Figure 8.6. In this case, the minimum most found was (0,-4.9) and was found 16.67% of the time. This is the first time a method found a local minimum more than any global minimum. The minimum least found was once again (0,0) at 3.47%. Therefore, we see that using this method there is a larger range in finding local minima than there is global minima.

We also compare the number of times the global and local minima were found in Figure 8.8. The global minimum was found 52.78% of the time and the local minimum was found 47.21% of the time. Using the systemized centroid method and large initial simplices, we once again find the global and local minimum a comparable amount, though the local minimum is found slightly more often. While this isn't ideal, we do recognize that the difference in global and local minima found here is smaller than the difference in global and local minima found in the systematic point method.

8.4 Conclusions

Using this function with multiple minima, we can further understand how the Nelder-Mead method works. We understand that despite a large number of minima, Nelder-Mead remains stable in the sense that it rarely stalls or fails to find a minimum given a simplex. We note than in all of the basin of attraction images, in this chapter and previous ones, there were only very few cases where an initial simplex was unable to find a minimum. This is an important distinction from other methods that become unable to find minima reliably when multiple minima are present.

Further, we once again note the differences between the systematic point method and the systemized centroid method are what we have consistently seen in previous chapters. The systematic point method does not find minima as we would expect it to based on the contour plot image but rather based on the two assigned points used to create the initial simplex. Even with several local and global minima, the systematic point method never showed preference for the global minima unless the initial simplex was sensitive towards them.

Meanwhile, the basin of attraction image for the systemized centroid method with a small initial simplex holds the same structure that we would expect it to based on the contour image. As we increase the size of the initial simplex, the structure of the basin of attraction image loosens and diverges from what we would expect to see. This method also finds global minima more reliably and frequently. Even with a function that had more local than global minima, the frequency of the global minima found was comparable to the frequency of local minima found.

Therefore, we maintain the idea that in most cases the systemized centroid method gives us the best understanding of the function from the basin of attraction image generated from the methods we worked with in this paper. It is not as sensitive to the initial simplex as the systematic point method and is easy to understand unlike the randomized centroid method. The systemized centroid method rarely stalls and fails to find a minimum and is the most precise at finding minima. It is the best, relatively, at finding global minima compared to local minima.

However, it is important to remember that this method seems to be the most effective of the ones we tested. There may be other ways to create the initial simplex or color code the basin attraction images to make them even more effective than the systemized centroid method. Also, the most effective method is relative to the function that we are studying. With other functions, we may find that one of the other methods creates a better basin of attraction image.

Bibliography

Abdelkader, A. “Maxima and Minima of Functions of Two Variables.” Maxima and Minima of Functions of Two Variables, 2021, https://www.analyzemath.com/calculus/multivariable/maxima_minima.html.

Fernández, Ander. How to Create Animations in Python, 4 Aug. 2021, https://anderfernandez.com/en/blog/how-to-create-animations-python/.

“File:Nelder mead2.Svg.” Wikimedia Commons, 20 Sept. 2016, https://commons.wikimedia.org/wiki/File:Nelder_Mead2.svg.

Levy, Adam B. Attraction in Numerical Minimization: Iteration Mappings, Attractors, and Basins of Attraction. Springer, 2018.

“Nelder-Mead Evolutionary Hybrid Algorithms (Artificial Intelligence).” Whatwhenhow RSS,http://what-when-how.com/artificial-intelligence/nelder-mead-evolutionary-hybrid-algorithms-artificial-intelligence/.

“Nelder–Mead Method.” Wikipedia, Wikimedia Foundation, 8 Apr. 2022, https://en.wikipedia.org/wiki/Nelder%E2%80%93Mead_method.

“Sudo Null - Latest It News.” SudoNull, 2017, https://sudonull.com/post/69185-Nelder-Mead-optimization-method-Python-implementation-example.