Chapter 5

Using a Local Embedding Method to find the dimension of a manifold and to develop an atlas of equations

Contents

5.1 Introduction: ......................................................... 39
5.2 The method of Local Embedding in a Dynamical System: . 40
5.3 Application of Local Embedding to a problem of Dynamics on a Mobius Strip: ........................................ 41
5.4 Role of Singular Value Decomposition in Local Embedding: 42
  5.4.1 Determination of the Local Dimension: ..................... 42
  5.4.2 Application of Non-linear Singular Value Decomposition for finding Local Dimension: ......................... 43
  5.4.3 Finding the Local Co-ordinates: .............................. 43
5.5 Low Dimensional Local Dynamics: ............................ 46
5.6 The Alignment of Local Patches: ................................. 49
5.7 An Application of Local Embedding in the case of Acoustic Signals: ......................................................... 53
5.8 Conclusion: .............................................................. 55

5.1 Introduction:

Earlier we have shown that if a dynamical system is embedded in a higher dimension it causes ambiguity in identifying equations from a given set of data. In the previous chapter, we have shown that in some specific cases, we can solve this problem by replacing the equation of the dynamical system by a certain type of equation. In reality this may not always be possible. In this chapter, we consider an alternative approach for finding equations from data in which we can avoid embedding completely.

As we have discussed in 2.3, often the minimum dimension that is required for embedding a system is higher than the dimension of the manifold in which its data
resides. For the harmonic oscillator, the dimension of the state space is 1, but we need a minimum of 2 dimensions for embedding. Many times, to identify equations from data we do not know what is the minimum dimension required for global embedding. Although for the Van Der Pol system 2 dimensions are enough for embedding, we embedded this system into 3 dimensions for finding the differential equation. We have shown that both the cases could lead to ambiguity. Moreover, in each of these cases data points are in high dimension (2 for circle), but the dimension of the manifold where this data resides is much low (1 for circle). Now, we examine if we can model any given set of data by a low dimensional system.

5.2 The method of Local Embedding in a Dynamical System:

Embedding can be classified as global and local. In global embedding we get a map or equation for the whole trajectory, but in local embedding we get a set of equations, each of which is valid only for a portion of the entire manifold. The validity of each equation is for a chart and since the charts overlap, we can follow the full set to calculate the predictions for the full manifold. In this method we can model our data by a low dimensional dynamics. If $N$-dimensional data resides in an $M$-dimensional manifold, where $N$ is less than $M$, we create an $N$-dimensional dynamics in each patch. So we can model our data by a low dimensional system.
5.3. Application of Local Embedding to a problem of Dynamics on a Mobius Strip:

the help of conventional geometric objects like curves and surfaces. So they are simple and easy to investigate. Often low-dimensional non-linear maps are the prototype models to study the emergence of complex behavior in dynamical systems. Low dimensional situations serve as excellent test cases for important dynamical phenomena like the dynamics of human voice or the dynamics of heart.

Finally, our physical world is three-dimensional and many of our ideas in topological dynamics are based on the intuition of our physical world. Hence it is more convenient for us to bring intuition and analogies from the physical world if we can model a problem in terms of a series of low dimensional systems.

5.3 Application of Local Embedding to a problem of Dynamics on a Mobius Strip:

We start with a simple example of the Mobius strip. Let us consider the parametric equation of the Mobius strip:

\[ X = 1 + v \cos(u/2)\cos(u) \]

\[ Y = 1 + v \cos(u/2)\sin(u) \]

\[ Z = v \sin(u/2) \] (5.1)

By varying \( u \) and \( v \), we generate a data set consisting of various values of \( X \), \( Y \) and \( Z \). Mobius strip is a 2-dimensional manifold, but it is embedded in 3 dimensions, so we get the data in 3 dimensions. \( u \) and \( v \) are parametrized by \( u_n = 2\pi 0.0001n \) and \( v_m = 0.0003m - 0.003 \), where \( n \) is from 0 to 21000 and \( m \) is from 0 to 20. We divide the whole manifold of the Mobius strip into overlapping local patches. We collect all the data points for \( n = 0 \)....14 and \( m = 0 \)....20 and denote it as the 0th patch. For \( n = 8 \)....22 and \( m = 0 \)....20 we get the 1st patch. In general, for \( n = 8w \)....8w + 14 and \( m = 0 \)....20 we get the \( w \)th patch. For each \( w \) we create a corresponding data matrix, say, \( H(w) \). \( H(w) \) has 3 columns, values of \( X \), \( Y \) and \( Z \). For each \( m \), we vary the value of \( n \) to get a string of the corresponding data points. Then we stack these strings one after another. Each data point has three components \( X \), \( Y \) and \( Z \). \( X \) component of the stacked strings forms the first column of \( H(w) \); \( Y \) and \( Z \) form the second and the third columns respectively. First let us consider the 0th patch and let us call the corresponding data matrix as \( H0 \).
5.4 Role of Singular Value Decomposition in Local Embedding:

By using Singular Value Decomposition (SVD), the 315 by 3 matrix $H_0$ can be decomposed as $H_0 = U W V^t$, where $U$ is a 315 by 3 matrix with the property $U^t U = I$, $W$ is a 3 by 3 diagonal matrix and $V$ is a 3 by 3 matrix with the property $V V^t = V^t V = I$. Diagonal elements of the matrix $W$ are the ordered non-negative square roots of the eigenvalues of $H_0^t H_0$. They are called the singular values of $H_0$. $V$ is a matrix made from a collection of the eigenvectors of $H_0^t H_0$ and $U$ is a sub-matrix of the collection of the eigenvectors of $H_0^t H_0$. This procedure is now well established and explained in reference [54]. Pictorially, this can be represented as shown in figure 5.3.

5.4.1 Determination of the Local Dimension:

To find the local dimension of a patch, we look at the number of the non-zero singular values. In the case of our numerical example, the singular values were: $(3.224, 0.0489, 0.004)$. We consider the ratio of the 1st singular value to the second singular value and the ratio of the 1st singular value to the 3rd singular value. The 1st to 2nd ratio is 100, while the 1st to 3rd ratio is much bigger. Hence we ignore the third one. So, even if we did not know that the Mobius strip data came from a two-dimensional manifold, we could have found it in this way.
5.4. Role of Singular Value Decomposition in Local Embedding:

5.4.2 Application of Non-linear Singular Value Decomposition for finding Local Dimension:

Non-linear singular value decomposition is an extension of the usual singular value decomposition [62]. The method is to augment the matrix with some additional non-linear columns derived from the initial vectors and extract the non-linear relationship by using SVD.

\[
H = \begin{pmatrix}
 x_0 & y_0 & z_0 \\
 x_1 & y_1 & z_1 \\
 x_2 & y_2 & z_2 \\
 \vdots & \vdots & \vdots \\
 x_n & y_n & z_n \\
\end{pmatrix}
\]

We create an extended \( H \) by augmenting the matrix \( H \) with some non-linear columns of \( X \), \( Y \) and \( Z \). Let us denote this extended \( H \) as \( EH \).

\[
EH = \begin{pmatrix}
 x_0 & y_0 & z_0 & x_0^2 & y_0^2 & z_0^2 & x_0 y_0 & y_0 z_0 & z_0 x_0 \\
 x_1 & y_1 & z_1 & x_1^2 & y_1^2 & z_1^2 & x_1 y_1 & y_1 z_1 & z_1 x_1 \\
 \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 x_n & y_n & z_n & x_n^2 & y_n^2 & z_n^2 & x_n y_n & y_n z_n & z_n x_n \\
\end{pmatrix}
\]

We see that the last singular value of \( EH \) is quite smaller than the first singular value of \( H \). The ratio of these two values is 0.0000206. In the case of the usual SVD, this ratio is 0.00124. Hence, non-linear SVD is a better way to find the local dimension of a manifold.

5.4.3 Finding the Local Co-ordinates:

Let us consider the following equation
$H0 = UWV^t \ldots \ldots (5.2)$. In this relation $V^t$ acts as an invertible, continuous and differentiable transformation between the data matrix $H0$ and the matrix $UW$. Let us see how this turns out to be the case, for a specific example. The 3rd singular value of $W$ is very small, so we can ignore the 3rd column of $UW$. Let us create a matrix with the first two columns of $UW$ and denote it as $sU$. $sU$ is a 100 by 2 matrix. Now let us consider the first two rows of $V^t$ and denote the sub-matrix as $sV^t$. Now, the above relationship becomes

$H0 = sUsV^t \ldots \ldots (5.3)$ Let us denote the columns of $sU$ as $(U0, U1)$. Let $sV^t = \begin{pmatrix} v_{0,0} & v_{0,1} \\ v_{1,0} & v_{1,1} \end{pmatrix}$ Hence we get,

$X = v_{0,0}U0 + v_{1,0}U1$

$Y = v_{0,1}U0 + v_{1,1}U1$

$Z = v_{0,2}U0 + v_{1,2}U1$

Each of $X$, $Y$ and $Z$ is a linear expansion of $U0$ and $U1$. So the transformation from $(X, Y, Z)$ to $(U0, U1)$ is continuous and differentiable. Now $V$ is an orthogonal matrix, i.e. $V^tV = I$. Therefore, $sV^tsV = I$. So $sV^t$ is an invertible transformation and

$H0sV = sU \ldots \ldots (5.4)$

From this relationship, we can express each of $U0$ and $U1$ as a linear combination of $X$, $Y$ and $Z$ in similar fashion. So the inverse transformation from $(U0, U1)$ to $(X, Y, Z)$ is also continuous and differentiable. This completes the outline of the proof that there is a diffeomorphism between $H0$ and $sU$. We have shown how we get a local diffeomorphism between the 3-dimensional data and the 2-dimensional
5.4. Role of Singular Value Decomposition in Local Embedding:

Figure 5.5: Local diffeomorphism between the 3 dim. data \((x,y,z)\) and the local co-ordinates \((U_0,U_1)\)
5.5 Low Dimensional Local Dynamics:

In section 5.4.3, we have discussed that there is a local diffeomorphism between the 3 dimensional data \((X, Y, Z)\) and the local co-ordinates \((U_0, U_1)\). We get a static relationship between \((X, Y, Z)\) and \((U_0, U_1)\) because of this local diffeomorphism. Equation (5.3) \(H_0 = sU_0S^V\) shows that there is a transformation from \((U_0, U_1)\) to \((X, Y, Z)\). Again from \((X, Y, Z)\) we can come back to \((U_0, U_1)\) by using the equation (5.4) \(H_0S^V = sU\). Now in each patch we find a local dynamics by using the local co-ordinates \((U_0, U_1)\). In the case of the Mobius strip, in each patch we get a 2 dim affine map of the form

\[
\begin{bmatrix}
U_{0n+1} \\
U_{1n+1}
\end{bmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{bmatrix} U_{0n} \\
U_{1n}
\end{bmatrix} + \begin{pmatrix} e \\
f \end{pmatrix}
\]

...(5.5) where \(U_0\) and \(U_1\) are the first two columns of \(U\) matrix. In this equation \(a, b, c, d, e\) and \(f\) are the unknowns. We create a matrix called \(D\), where \(D_{n,0} = U_{0n}\), \(D_{n,1} = U_{1n}\) and the third column is the constant vector \(1\). We create another matrix \(Pred\), which has two columns, \(Pred_{n,0} = U_{0n+1}, pred_{n,1} = U_{1n+1}\). Let \(C\) be the matrix of unknowns,

\[
C = \begin{pmatrix} a & c \\ b & d \\ e & f \end{pmatrix}
\]

From the above equation we get, \(DC = Pred\). By taking the generalized inverse of \(D\) we compute \(C\). For example, in the 0th patch \(H_0\), the patch values are \(a = 0.999997995, b = -0.0000096198, c = 0.0416904351, d = 0.99997518, e = -0.000000617, f = 0.0128276903\). This is the desired low dimensional dynamics for the patch \(H_0\). For each patch, we get a similar affine dynamics. In the 1st patch \(H_1\), \(a = 0.99999, b = -0.0000096197, c = 0.0416904351, d = 0.999997518, e = -0.000000617, f = 0.0128276937\). We consider the local co-ordinate \((U_0, U_1)\) of \(H_0\). We take the 0th value of \(U_0\) and \(U_1\) as the initial conditions and iterate it by the local affine map of \(H_0\). We get a good prediction. We see that we predict the dynamics in the green zone (in picture) quite well. When the overlapping area is reached, we go back to the data matrix \((x, y, z)\) from \((U_0, U_1)\) by using the relationship (5.3). Now a similar diffeomorphism exits between \(H_1\) and the corresponding \((U_0, U_1)\). By using this diffeomorphism again we come back to the \((U_0, U_1)\) which corresponds to the \(H_1\) patch (shown as a red rectangle in the picture). Now we consider the local co-ordinate \((U_0, U_1)\) of \(H_1\). We take the 0th value of \(U_0\) and \(U_1\) as the initial conditions and iterate it by the local affine map
Figure 5.6: left: predicted values of $U0$ in $H0$; right: actual values of $U0$ in $H0$
Chapter 5. Using a Local Embedding Method to find the dimension of a manifold and to develop an atlas of equations

Figure 5.7: left: predicted values of $U1$ in $H0$; right: actual values of $U1$ in $H0$

Figure 5.8: Local 2 dimensional dynamics
5.6. The Alignment of Local Patches: 

of $H1$. In this case too, we get a good prediction. Our numerical results show that we predict the dynamics in the red zone (in picture) quite well.

![Table of predicted values](image)

<table>
<thead>
<tr>
<th>$U_0$ in $H1$</th>
<th>$U_0$ in $H1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-0.0930495817</td>
</tr>
<tr>
<td>1</td>
<td>-0.0930490222</td>
</tr>
<tr>
<td>2</td>
<td>-0.0930485487</td>
</tr>
<tr>
<td>3</td>
<td>-0.0930481614</td>
</tr>
<tr>
<td>4</td>
<td>-0.0930478601</td>
</tr>
<tr>
<td>5</td>
<td>-0.0930476449</td>
</tr>
<tr>
<td>6</td>
<td>-0.0930475157</td>
</tr>
<tr>
<td>7</td>
<td>-0.0930474727</td>
</tr>
<tr>
<td>8</td>
<td>-0.0930475158</td>
</tr>
<tr>
<td>9</td>
<td>-0.0930476449</td>
</tr>
<tr>
<td>10</td>
<td>-0.0930478601</td>
</tr>
<tr>
<td>11</td>
<td>-0.0930481614</td>
</tr>
<tr>
<td>12</td>
<td>-0.0930485487</td>
</tr>
<tr>
<td>13</td>
<td>-0.0930490222</td>
</tr>
<tr>
<td>14</td>
<td>-0.0930495817</td>
</tr>
</tbody>
</table>

Figure 5.9: left: predicted values of $U0$ in $H1$; right: actual values of $U0$ in $H1$

5.6 The Alignment of Local Patches:

We have discussed how SVD gives us a local diffeomorphism between a local patch and a 2D rectangle. For each patch we get a corresponding 2D rectangle, but the patches are overlapping. We get rectangle 1 for patch 1 and rectangle 2 for patch 2. Now some points are common in both the patches, but we do not get any overlapping in the corresponding rectangles. So we keep the rectangle 1 as it is and give a transformation to the rectangle 2 so that the points corresponding to the overlapping part in the patches match each other. We continue doing this process until some patch (say Nth) overlaps with the 1st patch. Finally, we identify the overlapping points of the Nth rectangle with the overlapping points of the 1st rectangle.
Chapter 5. Using a Local Embedding Method to find the dimension of a manifold and to develop an atlas of equations

Figure 5.10: left: predicted values of $U_1$ in $H_1$; right: actual values of $U_1$ in $H_1$
Figure 5.11.a: Dynamics in two consecutive patches: We take the 0th value of $U_0$ and $U_1$ as the initial conditions and iterate it by the local affine map of $H_0$ to predict the dynamics in the green zone. When the overlapping area is reached, we go back to the data matrix $(x, y, z)$ from $(U_0, U_1)$ by using the relationship (5.3). Now a similar diffeomorphism exists between $H_1$ and the corresponding $(U_0, U_1)$. By using this diffeomorphism again we come back to the $(U_0, U_1)$ which corresponds to the $H_1$ patch (shown as a red rectangle in the picture). Now we consider the local co-ordinate $(U_0, U_1)$ of $H_1$. We take the 0th value of $U_0$ and $U_1$ as the initial conditions and iterate it by the local affine map of $H_1$ to predict the dynamics in the red zone.
Chapter 5. Using a Local Embedding Method to find the dimension of a manifold and to develop an atlas of equations

Figure 5.11.b: Alignment of two consecutive patches: We get the left rectangle for the left patch and the right rectangle for the right patch. Now some points are common in both the patches, but we do not get any overlapping in the corresponding rectangles. So we keep the left rectangle as it is and give a transformation to the right rectangle so that the points corresponding to the overlapping part in the patches match each other.

It is mentioned earlier that we get a local bijection between 3-dimensional data points and 2-dimensional rectangle by svd. But this bijection is not global. As for two overlapping patches, we get two non-intersecting rectangles. When we align rectangle 2 with rectangle 1 we create a global bijection between the 3 dimensional data and the 2 dimensional space. It is in fact a homeomorphism. So we actually create a covering space of the original manifold. In the case of the Mobius strip we get the following picture after aligning all of the 2D rectangles.

Figure 5.12: aligning all of the 2 dim. rectangles in a Mobius strip
5.7. An Application of Local Embedding in the case of Acoustic Signals:

We apply the above method to find the local dynamics of sound signal. We choose a sample of a female voice saying the vowel "e". The length of the sample is 327488. 44000 Hz is the sampling frequency of this signal. We get the following state space picture from the signal and we cover this picture by overlapping patches. Exploring the state space picture we choose 7 data points, say, \((P_0, Q_0, R_0), (P_1, Q_1, R_1), \ldots, (P_6, Q_6, R_6)\). For each \((P_k, Q_k, R_k)\) we collect all the data points whose Euclidean distance from \((P_k, Q_k, R_k)\) is less than 1.1 and we denote this neighbourhood or patch as \(N_k\). The 7 data points are chosen in such a way that any two consecutive patches overlap with each other.

Figure 5.14: State space picture of the vowel "e" signal
Chapter 5. Using a Local Embedding Method to find the dimension of a manifold and to develop an atlas of equations

Figure 5.15: covering the state space picture of vowel "e" by overlapping patches

For each patch, SVD was performed on the data inside that patch. One of the singular values always turned out to be quite small. For example, the singular values in the 0th patch are (84.41, 30.03, 0.78). The ratio of the last singular value to the first singular value is 0.00924. (if we use non-linear svd, we find that this ratio becomes 0.000036.) Thus we can safely conclude that the local dimension of each patch is 2. Following the same procedure as described in section 5.5, we arrive at the following 2D equation which predicts the dynamics locally. Our map in the 0th patch is

\[
\begin{bmatrix}
U_{0, n+1} \\
U_{1, n+1}
\end{bmatrix} = \begin{bmatrix}
a & b \\
c & d
\end{bmatrix} \times \begin{bmatrix}
U_{0, n} \\
U_{1, n}
\end{bmatrix} + \begin{bmatrix}
e \\
f
\end{bmatrix}
\]

where \(a = 0.9995, b = 0.0074, c = -0.194, d = 1.0019, e = 0\) and \(f = -0.0005\). Now if we take the 0th value of \(U0\) and \(U1\) as the initial conditions and iterate it by this affine map, we get a good prediction for \(U0\) and \(U1\).

By aligning the local patches together we get the following picture.
5.8 Conclusion:

In this chapter we have discussed the concept and technique of local embedding. We have improved the method of finding the local dimension of a manifold by using non-linear SVD. We have shown an application of this method in sound analysis, but in the case of ECG analysis this method does not work. Hence, for analyzing ECG data we have developed much more advanced techniques which are based on some topological ideas. In the rest of the thesis we show how topological concepts are used in local prediction of ECG and sound data.