# 中文：

假设一个UTS **X**嵌入到一个*m*维相空间中，那么我们可以得到一个*L* × *m*轨迹矩阵**Θ**。如果将**Θ**中的所有元素线性地映射到整数区间[0, 255]，则**Θ**也可以被视为分辨率为*L* × *m*的灰度图像。我们称**Θ**为轨迹矩阵图像(Trajectory Matrix Image, TMI）。在本文的剩余部分，除非另有规定，否则**Θ**表示TMI。TMI实际上是用二维图像表示的一维时间序列。这样，我们可以使用深度卷积神经网络自动地提取轨迹矩阵图像的特征。

图像**Θ**的尺寸(*L* × *m* = (*N* – (*m*-1)*τ*) × *m*)取决于*τ*和*m*。通常图像越大，DCNN的分类性能越好。因此，如何确定最佳的*τ*和*m*，使DCNN的分类性能达到最佳成为一个关键的问题。注意，互信息法和Cao方法等确定参数*τ*和*m*的常规方法不一定能使DCNN的分类性达到最佳。在本研究中，我们通过最大化TMI的分辨率来确定*τ*和*m*的最佳值。

**定理**：当*τ* = 1时， (或者), **Θ**的尺寸最大：

|  |  |
| --- | --- |
|  | (1) |

我们称对应的**Θ**为**最大轨迹矩阵图像**(maximum TMI, MTMI)，记为**Θ***max*。

**证明**：**Θ**的分辨率为*f*(*τ*, *m*) = *m* × (*N* – (*m*-1)*τ*)。为了获得**Θ***max*，可以通过最大化函数*f*来优化参数*τ*和*m*：

|  |  |
| --- | --- |
|  | (2) |

我们采用不等式约束的**拉格朗日乘子法**(Lagrange Multiplier)来求解*f*(*τ*, *m*)的极大值。为了简便，我们将原式改为：

|  |  |
| --- | --- |
|  | (3) |

则**拉格朗日函数**(Lagrange function)为：

|  |  |
| --- | --- |
|  | (0.4) |

其中*λ*为拉格朗日乘子。由不等式约束引入的**Karush-Kuhn-Tucker(KKT)**条件：

|  |  |
| --- | --- |
|  | (5) |

其中分别是变量*τ*、*m*关于函数的一阶偏导数：

|  |  |
| --- | --- |
|  | (6) |

我们需要对公式(6)分情况讨论。当时，，解得，代入得；当时，由于*m*为正整数而，因此只能取，此时，不合理，舍去。综上，由拉格朗日乘子法求得的*f*的极大值为*N*。

由于1 ≤ *m* ≤ *N*和1 ≤ *τ* < *N*，为了求得*f*的最大值，我们还需要考虑边界条件：

1. 当*m* = 1时，为常数，与无关；
2. 当*m* = *N*时，满足*τ*的约束条件，此时；
3. 当时，，在此条件下*m* = (*N* + 1) / 2取得最大值 (*N* + 1)2 / 4；
4. 当时，满足*m*的约束条件，此时。

由上我们可以得出(*N* + 1)2 / 4为*f*的最大值。因为代表嵌入维数，需为正整数。如果*N*为偶数，则*f*(1, *m*)的最大值等于*N*(*N* + 2) / 4；如果*N*为奇数，则*f*(1, *m*)的最大值等于(*N* + 1)2 / 4。在这个时候， 或者 .

图0.1显示了在区间1 ≤ *τ* < *N* 和 1 ≤ *m* ≤ *N* (以*N* = 11为例)中函数*f*(*τ*, *m*)的曲面图。注意，在绘制图0.1时，*τ*和*m*都被视为实数。从图0.1中可以看出，*f*(*τ*, *m*)在*τ*=1，*m*=6时达到最大值，这与定理1一致。

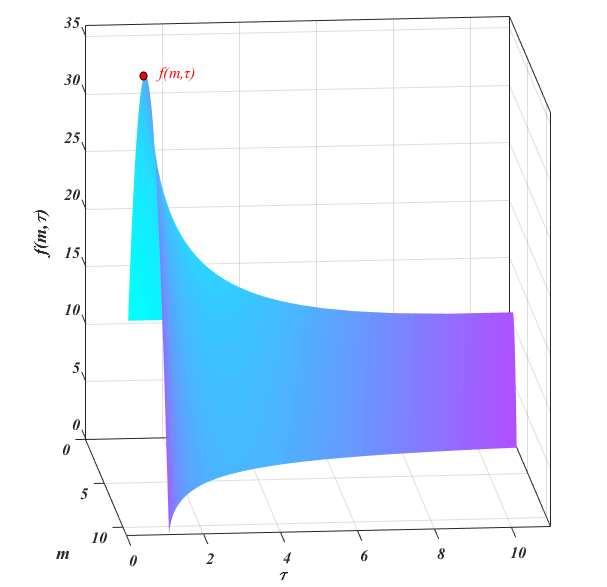


图0.1 在区间1 ≤ *τ* < *N* 和 1 ≤ *m* ≤ *N*中函数*f*(*τ*, *m*)的曲面图(以*N* = 11为例)

Fig. 0.1 Surface plot of the function *f*(*τ*, *m*) over the interval 1 ≤ *τ* < *N*和1 ≤ *m* ≤ *N*, taking *N* = 11 as an example

# 英文：

Assuming that a UTS **X** is embedded in *m*-dimensional phase space, then we get an *L × m* trajectory matrix **Θ**. If all elements of **Θ** are linearly mapped to integer interval [0, 255], **Θ** can also be regarded as a grayscale image with a resolution of *L × m*. We call **Θ** Trajectory Matrix Image (TMI). For the rest of this article, **Θ** means TMI unless otherwise specified. TMI is actually a one-dimensional time series represented by a two-dimensional image. In this way, we can use a deep convolutional neural network to automatically extract the features of the trajectory matrix image.

The size of **Θ** (*L* × *m* = (*N* – (*m*-1)*τ*) × *m*) depends on τ and *m*. Generally, the larger the image, the better the classification performance of DCNN. Therefore, how determining the best τ and *m* to achieve optimal classification performance of DCNN becomes a key problem. It should be noted that using conventional methods such as the mutual information method and Cao method to determine parameters τ and *m* may not make the classification of DCNN reach the best. In this study, we determine the best values of τ and *m* by maximizing the resolution of TMI.

**Theorem 1**: when the *τ* = 1, (or ) ), the **Θ** achieves the largest size:

|  |  |
| --- | --- |
|  | (1) |

the corresponding **Θ** is **MTMI** ( maximum TMI ), denoted as **Θ***max*.

**Proof 1**: the resolution of **Θ** is *f* (*τ*, *m*) = *m* × (*N* – (*m*-1)*τ*). To obtain **Θ***max*, the parameters *τ* and m can be optimized by maximizing the function *f*:

|  |  |
| --- | --- |
|  | (2) |

We use the Lagrange Multiplier method with inequality constraint to find the maximum value of *f(τ, m)*. For simplicity, we changed the original formula to:

|  |  |
| --- | --- |
|  | (3) |

Then **Lagrange function** is:

|  |  |
| --- | --- |
|  | (4) |

Where *λ* is the Lagrange multiplier. The **Karush-Kuhn-Tucker (KKT)** conditions introduced by the inequality constraints are as follows:

|  |  |
| --- | --- |
|  | (5) |

Where *Lm* and *Lτ* are the first partial derivatives of variables τ and m with respect to the function *L(τ, m, λ)* :

|  |  |
| --- | --- |
|  | (6) |

We need to discuss the situation of formula (6). When *g* < 0, *λ* = 0, solved as *m* = 1, *τ* = *N*. and *f* \* = – *N*; When *g* = 0, due to m being a positive integer and – λ ≤ 0, we can only take *m* = 1, where *N* = 1, which is unreasonable and should be rounded off. In conclusion, the local maximum value of *f* obtained by the Lagrange multiplier method is *N*.

Since *1 ≤ m ≤ N* and *1 ≤ τ < N*. To obtain the global maximum value of *f*, we also need to consider the boundary conditions:

1. When *m* = 1, *f (τ,* 1*)* = *N* is a constant, independent of *τ*;
2. When *m* = *N*, *τ* = 1 satisfies the constraint condition of *τ*, then *f (τ, N)* = *N*;
3. When *τ* = 1, *f* (1*, m*) = *- m2 + (N+*1*)m*, under this condition, we can obtain the maximum value *f* (1, *m*) = (*N*+1) 2 / 4 when *m* = (*N* + 1) / 2;
4. When τ=N, m =1 satisfies the constraint condition of m, then *f* (*N, m*) = *N*.

From the above, we can conclude that (*N*+1) 2 / 4 is the maximum value of *f* (*τ*, *m*). Because *m* represents the embedding dimension, it needs to be a positive integer. If *N* is an even number, then the maximum value of *f* (1, m) is equal to *N*(*N*+2) / 4; If N is odd, then the maximum value of *f* (1, m) is equal to (N+1) 2/4. At this time, or .

Fig. 1 shows the surface of the function *f* (*τ*, *m*) in the intervals 1 ≤ *τ* < *N* and 1 ≤ *m* ≤ *N* (take N = 11 as an example). Note that both *τ* and *m* are treated as real numbers when drawing Figure 3.2. It can also be seen from Figure 3.2 that *f* (*τ*, *m*) reaches its maximum value when *τ* = 1 and *m* = 6, which is consistent with Theorem 1.

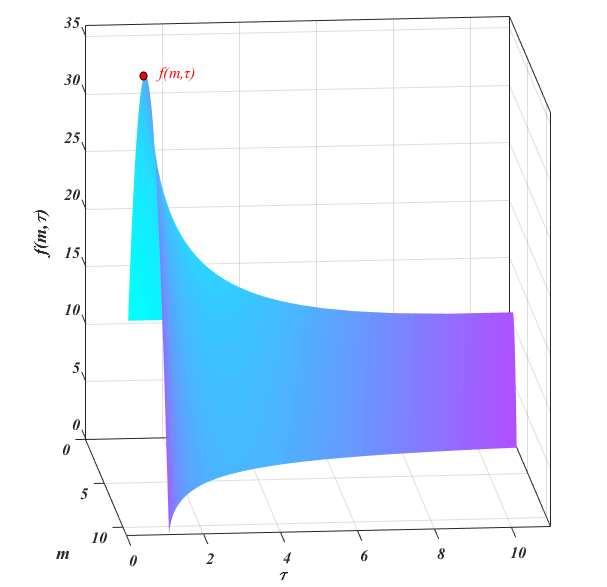


Fig. 1 Surface plot of the function *f*(*τ*, *m*) over the interval 1 ≤ *τ* < *N*和1 ≤ *m* ≤ *N*, taking *N* = 11 as an example