

MATLAB与C语言对比实例：导入txt数据

作者：凯鲁嘎吉 - 博客园 <http://www.cnblogs.com/kailugaji/>

1.MATLAB简介

MATLAB 是美国MathWorks公司出品的商业数学软件，用于算法开发、数据可视化、数据分析以及数值计算的高级技术计算语言和交互式环境。MATLAB是matrix&laboratory两个词的组合，意为矩阵工厂(矩阵实验室)，是由美国mathworks公司发布的主要面对科学计算、可视化以及交互式程序设计的高科技计算环境。它将数值分析、矩阵计算、科学数据可视化以及非线性动态系统的建模和仿真等诸多强大功能集成在一个易于使用的视窗环境中，为科学研究、工程设计以及必须进行有效数值计算的众多科学领域提供了一种全面的解决方案，并在很大程度上摆脱了传统非交互式程序设计语言(如C、Fortran)的编辑模式，代表了当今国际科学计算软件的先进水平。MATLAB可以进行矩阵运算、绘制函数和数据、实现算法、创建用户界面、连接其他编程语言的程序等，主要应用于工程计算、控制设计、信号处理与通讯、图像处理、信号检测、金融建模设计与分析等领域。MATLAB的基本数据单位是矩阵，简单易学，它的指令表达式与数学、工程中常用的形式十分相似，故用MATLAB来解算问题要比用C，FORTRAN等语言完成相同的事情简捷得多。在新的版本中也加入了对C，FORTRAN，C++，JAVA的支持。总之，matlab集成了很多数学计算功能，专注于数学计算领域，而C语言，是一个更大众化的程序平台。

2.实例分析：导入txt数据

1) 数据来源：UCI数据库中的数据

此电脑 > 文档 (E:) > 写在博客园				
	名称	修改日期	类型	大小
★	data	2018/5/6 17:18	文件夹	
	result	2018/5/6 17:43	文件夹	
	LoadData.m	2018/5/6 17:37	MATLAB Code	2 KB

此电脑 > 文档 (E:) > 写在博客园 > data				
名称	修改日期	类型	大小	
ecoli	2018/5/7 7:57	文件夹		
glass	2018/5/7 7:57	文件夹		
iris	2018/5/6 17:18	文件夹		

此电脑 > 文档 (E:) > 写在博客园 > data > iris				
名称	修改日期	类型	大小	
iris.txt	2008/8/6 14:38	文本文档	3 KB	
iris_data.txt	2008/12/4 15:44	文本文档	3 KB	
iris_id.txt	2008/12/4 15:45	文本文档	1 KB	
readme.txt	2010/10/9 15:48	文本文档	1 KB	

下面以iris数据为例

iris.txt中源数据如下:

```

5.1      3.5      1.4      0.2      0
4.9      3        1.4      0.2      0
4.7      3.2      1.3      0.2      0
4.6      3.1      1.5      0.2      0
5        3.6      1.4      0.2      0
5.4      3.9      1.7      0.4      0
4.6      3.4      1.4      0.3      0
5        3.4      1.5      0.2      0
4.4      2.9      1.4      0.2      0
4.9      3.1      1.5      0.1      0
5.4      3.7      1.5      0.2      0
4.8      3.4      1.6      0.2      0
4.8      3        1.4      0.1      0
4.3      3        1.1      0.1      0
5.8      4        1.2      0.2      0
5.7      4.4      1.5      0.4      0
5.4      3.9      1.3      0.4      0
5.1      3.5      1.4      0.3      0
5.7      3.8      1.7      0.3      0

```

5.1	3.8	1.5	0.3	0
5.4	3.4	1.7	0.2	0
5.1	3.7	1.5	0.4	0
4.6	3.6	1	0.2	0
5.1	3.3	1.7	0.5	0
4.8	3.4	1.9	0.2	0
5	3	1.6	0.2	0
5	3.4	1.6	0.4	0
5.2	3.5	1.5	0.2	0
5.2	3.4	1.4	0.2	0
4.7	3.2	1.6	0.2	0
4.8	3.1	1.6	0.2	0
5.4	3.4	1.5	0.4	0
5.2	4.1	1.5	0.1	0
5.5	4.2	1.4	0.2	0
4.9	3.1	1.5	0.2	0
5	3.2	1.2	0.2	0
5.5	3.5	1.3	0.2	0
4.9	3.6	1.4	0.1	0
4.4	3	1.3	0.2	0
5.1	3.4	1.5	0.2	0
5	3.5	1.3	0.3	0
4.5	2.3	1.3	0.3	0
4.4	3.2	1.3	0.2	0
5	3.5	1.6	0.6	0
5.1	3.8	1.9	0.4	0
4.8	3	1.4	0.3	0
5.1	3.8	1.6	0.2	0
4.6	3.2	1.4	0.2	0
5.3	3.7	1.5	0.2	0
5	3.3	1.4	0.2	0
7	3.2	4.7	1.4	1
6.4	3.2	4.5	1.5	1
6.9	3.1	4.9	1.5	1
5.5	2.3	4	1.3	1
6.5	2.8	4.6	1.5	1
5.7	2.8	4.5	1.3	1
6.3	3.3	4.7	1.6	1
4.9	2.4	3.3	1	1
6.6	2.9	4.6	1.3	1
5.2	2.7	3.9	1.4	1
5	2	3.5	1	1
5.9	3	4.2	1.5	1
6	2.2	4	1	1
6.1	2.9	4.7	1.4	1
5.6	2.9	3.6	1.3	1
6.7	3.1	4.4	1.4	1

5.6	3	4.5	1.5	1
5.8	2.7	4.1	1	1
6.2	2.2	4.5	1.5	1
5.6	2.5	3.9	1.1	1
5.9	3.2	4.8	1.8	1
6.1	2.8	4	1.3	1
6.3	2.5	4.9	1.5	1
6.1	2.8	4.7	1.2	1
6.4	2.9	4.3	1.3	1
6.6	3	4.4	1.4	1
6.8	2.8	4.8	1.4	1
6.7	3	5	1.7	1
6	2.9	4.5	1.5	1
5.7	2.6	3.5	1	1
5.5	2.4	3.8	1.1	1
5.5	2.4	3.7	1	1
5.8	2.7	3.9	1.2	1
6	2.7	5.1	1.6	1
5.4	3	4.5	1.5	1
6	3.4	4.5	1.6	1
6.7	3.1	4.7	1.5	1
6.3	2.3	4.4	1.3	1
5.6	3	4.1	1.3	1
5.5	2.5	4	1.3	1
5.5	2.6	4.4	1.2	1
6.1	3	4.6	1.4	1
5.8	2.6	4	1.2	1
5	2.3	3.3	1	1
5.6	2.7	4.2	1.3	1
5.7	3	4.2	1.2	1
5.7	2.9	4.2	1.3	1
6.2	2.9	4.3	1.3	1
5.1	2.5	3	1.1	1
5.7	2.8	4.1	1.3	1
6.3	3.3	6	2.5	2
5.8	2.7	5.1	1.9	2
7.1	3	5.9	2.1	2
6.3	2.9	5.6	1.8	2
6.5	3	5.8	2.2	2
7.6	3	6.6	2.1	2
4.9	2.5	4.5	1.7	2
7.3	2.9	6.3	1.8	2
6.7	2.5	5.8	1.8	2
7.2	3.6	6.1	2.5	2
6.5	3.2	5.1	2	2
6.4	2.7	5.3	1.9	2
6.8	3	5.5	2.1	2

5.7	2.5	5	2	2
5.8	2.8	5.1	2.4	2
6.4	3.2	5.3	2.3	2
6.5	3	5.5	1.8	2
7.7	3.8	6.7	2.2	2
7.7	2.6	6.9	2.3	2
6	2.2	5	1.5	2
6.9	3.2	5.7	2.3	2
5.6	2.8	4.9	2	2
7.7	2.8	6.7	2	2
6.3	2.7	4.9	1.8	2
6.7	3.3	5.7	2.1	2
7.2	3.2	6	1.8	2
6.2	2.8	4.8	1.8	2
6.1	3	4.9	1.8	2
6.4	2.8	5.6	2.1	2
7.2	3	5.8	1.6	2
7.4	2.8	6.1	1.9	2
7.9	3.8	6.4	2	2
6.4	2.8	5.6	2.2	2
6.3	2.8	5.1	1.5	2
6.1	2.6	5.6	1.4	2
7.7	3	6.1	2.3	2
6.3	3.4	5.6	2.4	2
6.4	3.1	5.5	1.8	2
6	3	4.8	1.8	2
6.9	3.1	5.4	2.1	2
6.7	3.1	5.6	2.4	2
6.9	3.1	5.1	2.3	2
5.8	2.7	5.1	1.9	2
6.8	3.2	5.9	2.3	2
6.7	3.3	5.7	2.5	2
6.7	3	5.2	2.3	2
6.3	2.5	5	1.9	2
6.5	3	5.2	2	2
6.2	3.4	5.4	2.3	2
5.9	3	5.1	1.8	2

前4列是属性取值，最后一列是分类结果

2) C语言：导入数据函数

```
//iris数据
#define data_num 150
#define data_dim 4
#define k_max 3
```

```

#define data_dir "iris"
#define dataset_name "iris"

typedef double elementtype;
elementtype data[data_num][data_dim];    //数据
int data_index[data_num];    //数据分类

//载入数据
void LoadData()
{
    int i, j;
    FILE *fp1, *fp2;
    char filename1[100];
    char filename2[100];
    double max, min;

    strcpy(filename1, "data//");
    strcat(filename1, data_dir);
    strcat(filename1, "//");
    strcat(filename1, dataset_name);
    strcat(filename1, "_data.txt");
    strcpy(filename2, "data//");
    strcat(filename2, data_dir);
    strcat(filename2, "//");
    strcat(filename2, dataset_name);
    strcat(filename2, "_id.txt");

    if((fp1=fopen(filename1, "r"))==NULL)
    {
        printf("can not open datafile!\n");
        exit(0);
    }
    if((fp2=fopen(filename2, "r"))==NULL)
    {
        printf("can not open datafile!\n");
        exit(0);
    }

    for(i=0; i<data_num; i++)
    {
        for(j=0; j<data_dim; j++)
            fscanf(fp1, "%lf", &data[i][j]);
        fscanf(fp2, "%d", &data_index[i]);
    }

    for(j=0; j<data_dim; j++)
    {

```

```

max=-999999999.0;
min=999999999.0;
for(i=0;i<data_num;i++)
{
    if (data[i][j]>max)
        max=data[i][j];
    if (data[i][j]<min)
        min=data[i][j];
}

for(i=0;i<data_num;i++)
    data[i][j]=(data[i][j]-min)/(max-min); //利用最大-最小值规范化到[0,1]区间
}

fclose(fp1);
fclose(fp2);
}

```

3) MATLAB语言：导入数据函数LoadData()

```

function [data,data_num,data_dim,data_index]=LoadData(data_dir,dataset_name)
%data是归一化属性取值，data_num是样例总数，data_dim是属性的个数，data_index是分类结果
format short
%载入真实数据
filename1='data/';
filename1=strcat(filename1,data_dir);
filename1=strcat(filename1,'/');
filename1=strcat(filename1,dataset_name);
filename1=strcat(filename1,'_data.txt');
filename2='data/';
filename2=strcat(filename2,data_dir);
filename2=strcat(filename2,'/');
filename2=strcat(filename2,dataset_name);
filename2=strcat(filename2,'_id.txt');

fp1=fopen(filename1,'r');
if fp1<=0
    print('can not open datafile!\n');
    exit(0);
else
    data=double(dlmread(filename1,'\t'));
end

fp2=fopen(filename2,'r');
if fp2<=0
    print('can not open datafile!\n');

```

```

        exit(0);
else
    data_index=dlmread(filename2, '\t');
end

[data_num,data_dim]=size(data);

for j=1:data_dim
    max=double(-999999999.0);
    min=double(999999999.0);
    for i=1:data_num
        if data(i,j)>max
            max=data(i,j);
        end
        if data(i,j)<min
            min=data(i,j);
        end
    end
    for i=1:data_num
        data(i,j)=double((data(i,j)-min)/(max-min));
    end
end
fclose(fp1);
fclose(fp2);

```

MATLAB结果:

```
>> [data,data_num,data_dim,data_index_first]=LoadData('iris','iris')
```

```
data =
```

0.2222	0.6250	0.0678	0.0417
0.1667	0.4167	0.0678	0.0417
0.1111	0.5000	0.0508	0.0417
0.0833	0.4583	0.0847	0.0417
0.1944	0.6667	0.0678	0.0417
0.3056	0.7917	0.1186	0.1250
0.0833	0.5833	0.0678	0.0833
0.1944	0.5833	0.0847	0.0417
0.0278	0.3750	0.0678	0.0417
0.1667	0.4583	0.0847	0
0.3056	0.7083	0.0847	0.0417
0.1389	0.5833	0.1017	0.0417
0.1389	0.4167	0.0678	0
0	0.4167	0.0169	0
0.4167	0.8333	0.0339	0.0417
0.3889	1.0000	0.0847	0.1250

0.3056	0.7917	0.0508	0.1250
0.2222	0.6250	0.0678	0.0833
0.3889	0.7500	0.1186	0.0833
0.2222	0.7500	0.0847	0.0833
0.3056	0.5833	0.1186	0.0417
0.2222	0.7083	0.0847	0.1250
0.0833	0.6667	0	0.0417
0.2222	0.5417	0.1186	0.1667
0.1389	0.5833	0.1525	0.0417
0.1944	0.4167	0.1017	0.0417
0.1944	0.5833	0.1017	0.1250
0.2500	0.6250	0.0847	0.0417
0.2500	0.5833	0.0678	0.0417
0.1111	0.5000	0.1017	0.0417
0.1389	0.4583	0.1017	0.0417
0.3056	0.5833	0.0847	0.1250
0.2500	0.8750	0.0847	0
0.3333	0.9167	0.0678	0.0417
0.1667	0.4583	0.0847	0.0417
0.1944	0.5000	0.0339	0.0417
0.3333	0.6250	0.0508	0.0417
0.1667	0.6667	0.0678	0
0.0278	0.4167	0.0508	0.0417
0.2222	0.5833	0.0847	0.0417
0.1944	0.6250	0.0508	0.0833
0.0556	0.1250	0.0508	0.0833
0.0278	0.5000	0.0508	0.0417
0.1944	0.6250	0.1017	0.2083
0.2222	0.7500	0.1525	0.1250
0.1389	0.4167	0.0678	0.0833
0.2222	0.7500	0.1017	0.0417
0.0833	0.5000	0.0678	0.0417
0.2778	0.7083	0.0847	0.0417
0.1944	0.5417	0.0678	0.0417
0.7500	0.5000	0.6271	0.5417
0.5833	0.5000	0.5932	0.5833
0.7222	0.4583	0.6610	0.5833
0.3333	0.1250	0.5085	0.5000
0.6111	0.3333	0.6102	0.5833
0.3889	0.3333	0.5932	0.5000
0.5556	0.5417	0.6271	0.6250
0.1667	0.1667	0.3898	0.3750
0.6389	0.3750	0.6102	0.5000
0.2500	0.2917	0.4915	0.5417
0.1944	0	0.4237	0.3750
0.4444	0.4167	0.5424	0.5833
0.4722	0.0833	0.5085	0.3750

0.5000	0.3750	0.6271	0.5417
0.3611	0.3750	0.4407	0.5000
0.6667	0.4583	0.5763	0.5417
0.3611	0.4167	0.5932	0.5833
0.4167	0.2917	0.5254	0.3750
0.5278	0.0833	0.5932	0.5833
0.3611	0.2083	0.4915	0.4167
0.4444	0.5000	0.6441	0.7083
0.5000	0.3333	0.5085	0.5000
0.5556	0.2083	0.6610	0.5833
0.5000	0.3333	0.6271	0.4583
0.5833	0.3750	0.5593	0.5000
0.6389	0.4167	0.5763	0.5417
0.6944	0.3333	0.6441	0.5417
0.6667	0.4167	0.6780	0.6667
0.4722	0.3750	0.5932	0.5833
0.3889	0.2500	0.4237	0.3750
0.3333	0.1667	0.4746	0.4167
0.3333	0.1667	0.4576	0.3750
0.4167	0.2917	0.4915	0.4583
0.4722	0.2917	0.6949	0.6250
0.3056	0.4167	0.5932	0.5833
0.4722	0.5833	0.5932	0.6250
0.6667	0.4583	0.6271	0.5833
0.5556	0.1250	0.5763	0.5000
0.3611	0.4167	0.5254	0.5000
0.3333	0.2083	0.5085	0.5000
0.3333	0.2500	0.5763	0.4583
0.5000	0.4167	0.6102	0.5417
0.4167	0.2500	0.5085	0.4583
0.1944	0.1250	0.3898	0.3750
0.3611	0.2917	0.5424	0.5000
0.3889	0.4167	0.5424	0.4583
0.3889	0.3750	0.5424	0.5000
0.5278	0.3750	0.5593	0.5000
0.2222	0.2083	0.3390	0.4167
0.3889	0.3333	0.5254	0.5000
0.5556	0.5417	0.8475	1.0000
0.4167	0.2917	0.6949	0.7500
0.7778	0.4167	0.8305	0.8333
0.5556	0.3750	0.7797	0.7083
0.6111	0.4167	0.8136	0.8750
0.9167	0.4167	0.9492	0.8333
0.1667	0.2083	0.5932	0.6667
0.8333	0.3750	0.8983	0.7083
0.6667	0.2083	0.8136	0.7083
0.8056	0.6667	0.8644	1.0000

0.6111	0.5000	0.6949	0.7917
0.5833	0.2917	0.7288	0.7500
0.6944	0.4167	0.7627	0.8333
0.3889	0.2083	0.6780	0.7917
0.4167	0.3333	0.6949	0.9583
0.5833	0.5000	0.7288	0.9167
0.6111	0.4167	0.7627	0.7083
0.9444	0.7500	0.9661	0.8750
0.9444	0.2500	1.0000	0.9167
0.4722	0.0833	0.6780	0.5833
0.7222	0.5000	0.7966	0.9167
0.3611	0.3333	0.6610	0.7917
0.9444	0.3333	0.9661	0.7917
0.5556	0.2917	0.6610	0.7083
0.6667	0.5417	0.7966	0.8333
0.8056	0.5000	0.8475	0.7083
0.5278	0.3333	0.6441	0.7083
0.5000	0.4167	0.6610	0.7083
0.5833	0.3333	0.7797	0.8333
0.8056	0.4167	0.8136	0.6250
0.8611	0.3333	0.8644	0.7500
1.0000	0.7500	0.9153	0.7917
0.5833	0.3333	0.7797	0.8750
0.5556	0.3333	0.6949	0.5833
0.5000	0.2500	0.7797	0.5417
0.9444	0.4167	0.8644	0.9167
0.5556	0.5833	0.7797	0.9583
0.5833	0.4583	0.7627	0.7083
0.4722	0.4167	0.6441	0.7083
0.7222	0.4583	0.7458	0.8333
0.6667	0.4583	0.7797	0.9583
0.7222	0.4583	0.6949	0.9167
0.4167	0.2917	0.6949	0.7500
0.6944	0.5000	0.8305	0.9167
0.6667	0.5417	0.7966	1.0000
0.6667	0.4167	0.7119	0.9167
0.5556	0.2083	0.6780	0.7500
0.6111	0.4167	0.7119	0.7917
0.5278	0.5833	0.7458	0.9167
0.4444	0.4167	0.6949	0.7083

data_num =

```
data_dim =
```

4

```
data_index_first =
```

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

O O O O O O O O O O

[illegible]

[illegible]

3. 总结

MATLAB的底层是C写的，C的效率比MatLab要高，但MatLab语法简单，但执行效率不高。MATLAB是用于特定的方面的，比如说矩阵运算方面、DSP就很有优势，而C语言不同，很多软件的底层都是C编写的。MATLAB与C语言有很多相似之处，学习了C语言后，MATLAB就会相对易学。

如函数有不足之处，望园友多多指教。