

# MATLAB实例：求相关系数、绘制热图并找到强相关对

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

用MATLAB编程，求给定数据不同维度之间的相关系数，并绘制热图，保存图片，找到强相关的维度对。第3部分还给定相关系数，并自己DIY热力图颜色，给出实例。

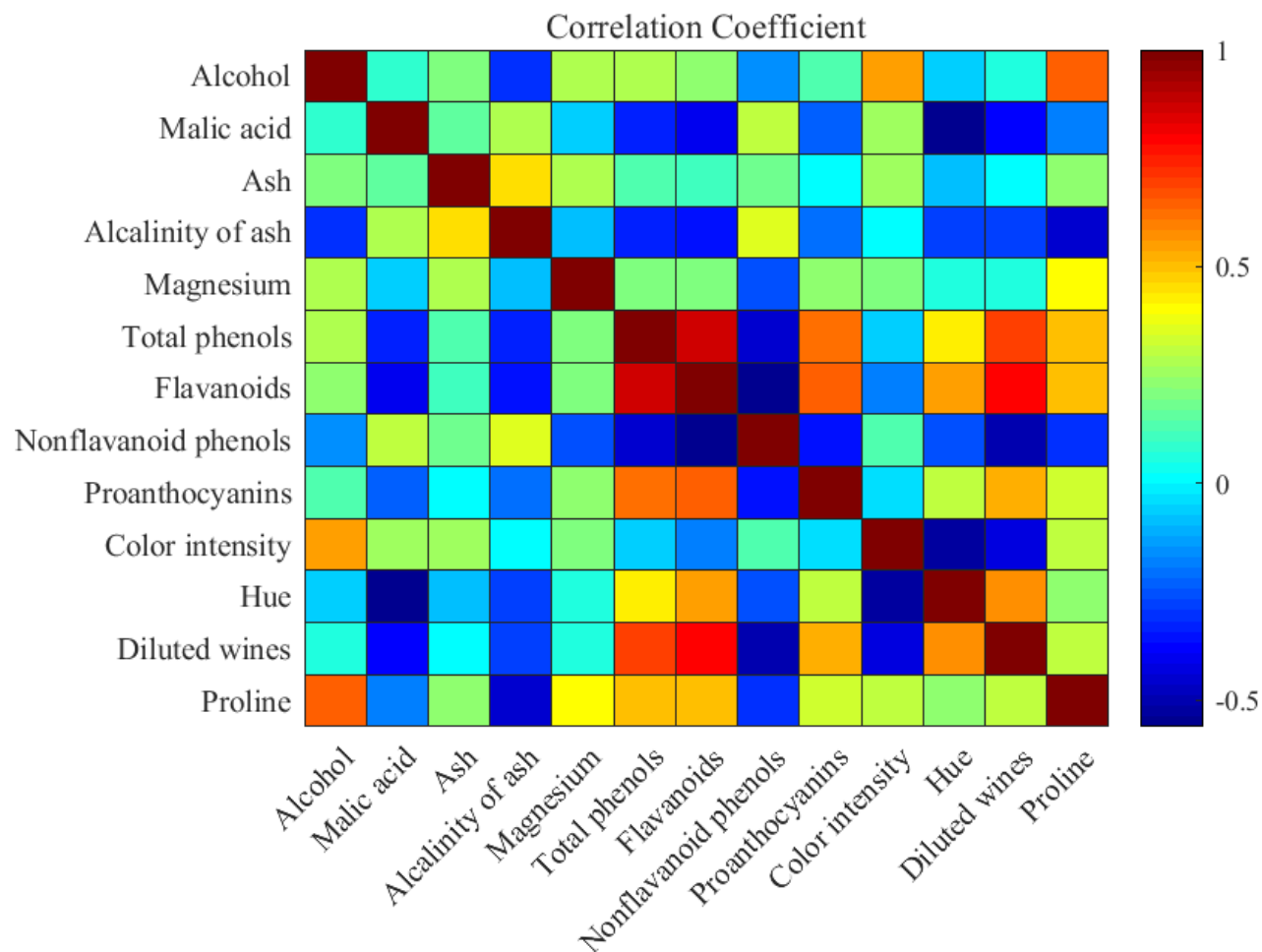
数据集来自UCI中的wine: <https://archive.ics.uci.edu/ml/datasets/Wine>

## 1. demo.m

```
clear
clc
etea=0.8; %阈值
% 加载数据
data_load=dlmread('E:\scanplot\wine.data');
data=data_load(:,2:14);
[N, D]=size(data);
% 求维度之间的相关系数
rho = corr(data, 'type','pearson');
% 绘制热图
string_name={'Alcohol','Malic acid','Ash','Alcalinity of ash','Magnesium','Total phenols','Flavanoids','Nonflavanoid phenols','Proanthocyanins','Color intensity','Hue','Diluted wines','Proline'};
xvalues = string_name;
yvalues = string_name;
h = heatmap(xvalues,yvalues,rho, 'FontSize',10, 'FontName','Times New Roman');
h.Title = 'Correlation Coefficient';
colormap(jet)
saveas(gcf,sprintf('wine相关系数热图.jpg'),'bmp'); %保存图片
% 绝对值
rho=abs(rho);
rho_l=rho.*tril(ones(D,D),-1); %下三角
[row, col]=find(rho_l>etea); %找>etea的两个维度
[Num, ~]=size(row);
% A:存放相关系数>etea的两个维度及相关系数值
A=zeros(Num, 3);
for i=1:Num
    A(i,:)=[row(i), col(i), rho_l(row(i), col(i))];
    fprintf('强线性相关的两个维度是：第%d个维度：%s与第%d个维度：%s，两者的相关系数为：%f\n', row(i), string_name{row(i)}, col(i), string_name{col(i)}, rho_l(row(i), col(i)));
end
```

## 2. 结果

```
>> demo
强线性相关的两个维度是：第7个维度：Flavanoids与第6个维度：Total phenols，两者的相关系数为：0.864564
```



### 3. 给定相关系数并自定义热力图颜色

```
% 自己给定相关系数，自己定义热力图颜色
% 作者：凯鲁嘎吉 - 博客园 http://www.cnblogs.com/kailugaji/
clear
clc
figure(1)
rho=[1.00    0.16    0.29    0.05    0.34    0.41    0.29    0.22    0.25    0.56    0.37    0.19    0.44
0.16    1.00    0.44    0.29    0.13    0.12    0.19    0.01    0.26    0.07    0.00    0.22    0.05
0.29    0.44    1.00    0.08    0.32    0.35    0.36    0.20    0.02    0.27    0.28    0.44    0.52
0.05    0.29    0.08    1.00    0.21    0.20    0.26    0.24    0.20    0.06    0.07    0.39    0.21
0.34    0.13    0.32    0.21    1.00    0.86    0.45    0.61    0.06    0.43    0.70    0.50    0.72
0.41    0.12    0.35    0.20    0.86    1.00    0.54    0.65    0.17    0.54    0.79    0.49    0.85
0.29    0.19    0.36    0.26    0.45    0.54    1.00    0.37    0.14    0.26    0.50    0.31    0.49
0.22    0.01    0.20    0.24    0.61    0.65    0.37    1.00    0.03    0.30    0.52    0.33    0.50
0.25    0.26    0.02    0.20    0.06    0.17    0.14    0.03    1.00    0.52    0.43    0.32    0.27]
```

```

0.56    0.07    0.27    0.06    0.43    0.54    0.26    0.30    0.52    1.00    0.57    0.24    0.62
0.37    0.00    0.28    0.07    0.70    0.79    0.50    0.52    0.43    0.57    1.00    0.31    0.79
0.19    0.22    0.44    0.39    0.50    0.49    0.31    0.33    0.32    0.24    0.31    1.00    0.63
0.44    0.05    0.52    0.21    0.72    0.85    0.49    0.50    0.27    0.62    0.79    0.63    1.00
]; %自己给定相关系数
% 绘制热图
string_name={'Alcohol','Malic acid','Ash','Alcalinity of ash','Magnesium','Total phenols','Flavanoids','Nonflavanoid phenols','Proanthocyanins','Color intensity','Hue','Diluted wines','Proline'};
xvalues = string_name;
yvalues = string_name;
hl = heatmap(xvalues,yvalues, rho, 'FontSize',10, 'FontName','Times New Roman');
hl.Title = 'Correlation Coefficient';
% hl.ColorbarVisible = 'off';
map = [1 1 1; 1 1 0; 0.5 1 0.4; 0.2 0.85 0.2; 0.4 0.7 1; 0.2 0.5 0.8]; % 自己定义颜色
colormap(map)
% saveas(gcf,sprintf('wine相关系数热图_自定义.jpg'),'bmp'); %保存图片

```

