

字符数组

讲义和代码可以看配套的课程第一个视频下载

配套课程b站：《MATLAB教程新手入门篇（数学建模清风主讲，适合零基础同学观看）》
数学建模清风老师版权所有

ASCII和Unicode编码

ASCII编码总共定义了128个字符，这128个字符的十进制编号为0到127。

在ASCII编码表中，常用字符的编码（十进制）如下所示：
数字：48至57表示 '0' 至 '9'
大写字母：65至90表示 'A' 至 'Z'
小写字母：97至122表示 'a' 至 'z'
编码10表示换行符；编码32表示空格

Unicode编码为每种语言中的每个字符设定了统一且唯一的编码，以满足跨语言、跨平台的要求

单个字符

- 字符(char)是文本的最基本单元。在MATLAB中，单个字符使用单引号引起来，例如：'a'、'0'、'我'。
- double函数可以获取字符对应的Unicode编码(十进制)。例如double('a')返回97
- char函数可以将Unicode编码(十进制)转换为对应的字符。例如char(97)返回'a'
- 当对字符进行算术运算时，MATLAB会自动将其转换为对应的十进制编码。例如'a' - 'A'等于32；'A' + 32等于97

将多个字符放到同一个单引号中就能创建一个字符向量，例如：'abc'、'001'、'我喜欢你'。

字符向量也分为行字符向量和列字符向量。行字符向量的使用频率更高！在不引起混淆的情况下，我们通常将行字符向量简称为字符向量

我们可以对字符向量使用double函数，这样可以得到该字符向量中各字符元素对应的Unicode编码(十进制)；类似的，我们也可以将一个数值向量使用char函数转换为字符向量。

我们可以对字符向量使用 double 函数，这样可以得到该字符向量中各字符元素对应的Unicode 编码(十进制)，返回的是一个数值向量。

类似的，我们也可以将一个数值向量使用 char 函数转换为字符向量。

字符向量也能进行算术运算和关系运算

字符向量也能通过索引的方式进行引用、修改和删除

我们还可以对字符向量进行拼接、重复、排序、翻转等操作

字符向量

(1) 字符向量中有单引号怎么办

使用两个连续的单引号来表示它

c2 = 'I'm Li Hua.'
c3 = 'a''b'

(2) 字符向量中如何添加换行符

换行符对应的Unicode编码为10，因此我们可以使用char(10)得到一个换行符

MATLAB中提供了创建换行符的函数newline，它和char(10)完全相同

['这是第一行',newline,'第二行喔']
ans =
'这是第一行
第二行喔'

(3) 计算字符向量中字符的数量

我们可以使用length 函数或 numel 函数计算字符向量中字符的数量，如果使用size函数，那么将同时返回字符向量的行数和列数

c = 'abcdefg';
numel(c)
length(c)
size(c)

(4) 比较两个字符向量是否相同

第一种：关系运算符==

c1 = 'pool';
c2 = 'moon';
c1 == c2
c3 = '你说你喜欢我，希望你没说谎';
ind = (c3 == '你')
num = sum(ind) % c3 中 '你' 出现的次数
c4 = 'abc';
c5 = ['a';'b';'c']; % 列字符向量
c4 == c5
c6 = 'abc';
c7 = 'abcd';
c6 == c7

第二种：比较两个字符向量是否完全相同

strcmp 比较字符向量是否相同（区分大小写）
strcmpi 比较字符向量是否相同（不区分大小写）
strncmp 比较字符向量的前 n 个字符是否相同（区分大小写）
strncmpi 比较字符向量的前 n 个字符是否相同（不区分大小写）

% 比较整个字符向量，区分大小写
strcmp('abc','Abc')
% 比较整个字符向量，不区分大小写
strcmpi('abc','Abc')
% 比较前 3 个字符，区分大小写
strncmp('wxyz','wxy012',3)
% 比较前 3 个字符，不区分大小写
strncmpi('wxyz','wxy012',3)

练习题中提到的新函数或者新的知识点

newStr = upper(str) 将 str 中的所有小写字母转换为相应的大写字母并保留所有其他字符不变；
相应地，内置函数lower可以将所有的大写字母转换为相应的小写字母，并保留所有其他字符不变。

...(三个点/省略号) 表示行的延续。有时候一行代码很长，为了便于阅读和理解，我们可以使用 ... 将这行代码分割到多行。

num2str函数，它能够将数字转换成字符的形式

num2str(520)
num2str(pi) % 将圆周率π转换为字符向量
num2str(pi,3) % 指定有效数字的最大位数为 3
num2str(pi,10) % 指定有效数字的最大位数为 10
num2str(1:10) % 将行向量转换为对应的字符向量
% MATLAB 会自动添加空格来隔开行向量的元素
num2str([1;2;3]) % 将列向量进行转换
x = [3 4;
1 10;
7 8];
num2str(x) % 将数值矩阵转换为对应的字符矩阵

dec2base函数可以将十进制(decimalism)整数转换为其 n 进制表示形式：dec2base(D, n) 返回十进制整数 D 的 n 进制表示形式。

dec2base(38,2) % 将38转换成二进制(binary)
% 等价于dec2bin(38), dec2bin函数可以将十进制转换为二进制
% 返回结果：'100110'，和我们自己写的代码结果一致

如果我们想在同一个变量中表示多段文本，那么就可以使用字符矩阵。

上面两种写法生成的字符矩阵c1和c2完全一样

c = ['good';
'great']

思考：如何在字符矩阵中表示good和great这两个单词？

MATLAB报错的原因是要串联的数组的维度不一致，这是因为字符向量'good'中有 4 个字符，而字符向量'great'中却有五个字符，因此不满足纵向拼接的条件。那么应该如何解决呢？我们可以在 good 后面加上一个空格：

c = ['good '
'great']

blanks(n)可以生成n个空格构成的字符向量

C = char(A1,A2,...,An) 将输入的数组 A1,A2,...,An（通常是字符向量、字符矩阵，也能支持数值向量、数值矩阵等）转换为单个字符数组，输入的A1,A2,...,An会变成 C 中的行。

char 函数会自动在每一行最后使用空格填充来保证每行字符数量相同

例如，如果我们有几个不同长度的字符向量，例如 'good'、' '（空字符向量）、'great' 和 'perfect'，并希望将它们组合成一个字符矩阵，我们可以使用以下命令：

char('good',' ','great','perfect')

strvcat函数也可以用于纵向拼接字符数组，它和char函数的区别在于：拼接过程中，char函数不会忽略空字符向量（会使用空格填充这行），而strvcat会忽略空字符向量。

deblank函数只能删除字符向量末尾的空白字符，不会删除开头的空白字符；

拓展一：deblank、strtrim 和 strip 函数
这三个函数用来删除字符向量开头或末尾的空白字符

strtrim函数会同时删除字符向量开头和末尾的空白字符；

strip函数则提供了更多的选项，例如可以指定删除哪一侧的空白字符（左侧 'left'、右侧 'right'、双侧 'both'），甚至还能自己指定要删除的字符，即不一定是空白字符。

strjust函数用于调整字符数组中文本的对齐方式，它通过调整每一行字符向量前后空格的位置来实现左对齐(left)、右对齐(right, 默认值)和居中对齐(center)三种格式。

拓展二：strjust函数

c = char('good','great','wonderful');
disp(c)
c1 = strjust(c); % 右对齐格式
disp(c1) % 也可以写完整 strjust(c,'right')
c2 = strjust(c,'center'); % 居中对齐格式
disp(c2)