

第6章 MATLAB统计函数

MATLAB统计分析

- 统计分析函数
 - `max(x)`—找出`x`的最大值
 - `max(x,y)`—找出`x`和`y`的最大值
 - `[y,i]=max(x)`—找出`x`的最大值`y`及位置`i`
 - `cumsum(x)`—计算数组`x`的累加值
 - `std(x)`—数组`x`的元素的方差
 - `cumprod(x)`—计算数组`x`的累加连乘值
 - `mean(x)`—计算`x`的平均值
 - `median(x)`—计算`x`的中位数
 - `sum(x)`—计算数值`x`的和

6.1 数据统计处理

6.1.1 最大值和最小值

MATLAB提供的求数据序列的最大值和最小值的函数分别为**max**和**min**，两个函数的调用格式和操作过程类似。

1. 求向量的最大值和最小值

求一个向量**X**的最大值的函数有两种调用格式，分别是：

(1) y=max(X)：返回向量**X**的最大值存入**y**，如果**X**中包含复数元素，则按模取最大值。

(2) [y,I]=max(X): 返回向量X的最大值存入y，最大值的序号存入I，如果X中包含复数元素，则按模取最大值。

求向量X的最小值的函数是**min(X)**，用法和**max(X)**完全相同。

例6-1 求向量x的最大值。

命令如下：

```
x=[-63,72,9,16,23,67];
```

```
y=max(x)           %求向量x中的最大值
```

```
[y,I]=max(x)       %求向量x中的最大值及其该元素的位置
```

2. 求矩阵的最大值和最小值

求矩阵A的最大值的函数有3种调用格式，分别是：

- (1) **max(A)**: 返回一个行向量，向量的第i个元素是矩阵A的第i列上的最大值。
- (2) **[Y,U]=max(A)**: 返回行向量Y和U，Y向量记录A的每列的最大值，U向量记录每列最大值的行号。

(3) **max(A,[],dim)**: **dim**取1或2。**dim**取1时, 该函数和**max(A)**完全相同; **dim**取2时, 该函数返回一个列向量, 其第*i*个元素是A矩阵的第*i*行上的最大值。

求最小值的函数是**min**, 其用法和**max**完全相同。

例6-2 分别求 3×6 矩阵x中各列和各行元素中的最大值, 并求整个矩阵的最大值和最小值。

3. 两个向量或矩阵对应元素的比较

函数**max**和**min**还能对两个同型的向量或矩阵进行比较，调用格式为：

(1) **U=max(A,B)**: **A,B**是两个同型的向量或矩阵，结果**U**是与**A,B**同型的向量或矩阵，**U**的每个元素等于**A,B**对应元素的较大者。

(2) **U=max(A,n)**: **n**是一个标量，结果**U**是与**A**同型的向量或矩阵，**U**的每个元素等于**A**对应元素和**n**中的较大者。

min函数的用法和**max**完全相同。

注意： **max(A,[],dim)**和 **max(A,dim)**的差异。

例6-3 求两个 2×3 矩阵**x**, **y**所有同一位置上的较大元素构成的新矩阵**p**。

6.1.2 求和与求积

数据序列求和与求积的函数是**sum**和**prod**，其使用方法类似。设**X**是一个向量，**A**是一个矩阵，

函数的调用格式为：

sum(X)：返回向量**X**各元素的和。

prod(X)：返回向量**X**各元素的乘积。

sum(A)：返回一个行向量，其第*i*个元素是**A**的第*i*列的元素和。

prod(A)：返回一个行向量，其第*i*个元素是**A**的第*i*列的元素乘积。

sum(A,dim): 当dim为1时，该函数等同于sum(A)；当dim为2时，返回一个列向量，其第i个元素是A的第i行的各元素之和。

prod(A,dim): 当dim为1时，该函数等同于prod(A)；当dim为2时，返回一个列向量，其第i个元素是A的第i行的各元素乘积。

例6-6 求矩阵A的每行元素的乘积和全部元素的乘积。

6.1.3 平均值和中值

mean:求数据序列平均值的函数;

median:求数据序列中值的函数。

两个函数的调用格式为:

mean(X): 返回向量X的算术平均值。

median(X): 返回向量X的中值。

mean(A): 返回一个行向量，其第i个元素是A的第i列的算术平均值。

median(A): 返回一个行向量，其第i个元素是A的第i列的中值。

mean(A,dim):

当**dim**为1时，该函数等同于**mean(A)**;

当**dim**为2时，返回一个列向量，其第**i**个元素是**A**的第**i**行的算术平均值。

median(A,dim):

当**dim**为1时，该函数等同于**median(A)**;

当**dim**为2时，返回一个列向量，其第**i**个元素是**A**的第**i**行的中值。

例6-6 分别求向量**x**与**y**的平均值和中值。

6.1.4 累加和与累乘积

在MATLAB中，使用**cumsum**和**cumprod**函数能方便求得向量和矩阵元素的累加和与累乘积向量，函数的调用格式为：

cumsum(X): 返回向量X累加和向量。

cumprod(X): 返回向量X累乘积向量。

cumsum(A): 返回一个矩阵，其第i列是A的第i列的累加和向量。

cumprod(A): 返回一个矩阵，其第i列是A的第i列的累乘积向量。

cumsum(A,dim):

当**dim**为1时，该函数等同于**cumsum(A)**;

当**dim**为2时，返回一个矩阵，其第**i**行是**A**的第**i**行的累加和向量。

cumprod(A,dim):

当**dim**为1时，该函数等同于**cumprod(A)**;

当**dim**为2时，返回一个向量，其第**i**行是**A**的第**i**行的累乘积向量。

例6-6 求**s**矩阵的累加和。

6.1.6 标准方差与相关系数

1. 求标准方差

std: 计算数据序列的标准方差的函数。

对于向量**X**, **std(X)**返回一个标准方差。

对于矩阵**A**, **std(A)**返回一个行向量, 它的各个元素便是矩阵**A**各列的标准方差。

std函数的一般调用格式为:

$$Y = \text{std}(A, \text{flag}, \text{dim})$$

flag取0或1，

当**flag**=0时，按**σ1**所列公式计算标准方差；

当**flag**=1时，按**σ2**所列公式计算标准方差。

缺省**flag**=0，**dim**=1。

其中**dim**取1或2

当**dim**=1时，求各列元素的标准方差；

当**dim**=2时，则求各行元素的标准方差。

例6-7 对二维矩阵**x**，从不同维方向求出其标准方差。

2、相关系数

MATLAB提供了corrcoef函数，可以求出数据的相关系数矩阵。

corrcoef函数的调用格式为： `corrcoef(X)`

返回从矩阵X形成的一个相关系数矩阵。此相关系数矩阵的大小与矩阵X一样。它把矩阵X的每列作为一个变量，然后求它们的相关系数。

执行**`A=corrcoef(X)`**后，**A**中每个值的所在行**a**和列**b**，反映的是原矩阵**X**中相应的第**a**个列向量和第**b**个列向量的相似程度（即相关系数,**corr**）。

3、互相关函数

`C=xcorr (A, B)`

例6-8 生成满足正态分布的**10000×6**随机矩阵，然后求各列元素的均值和标准方差，再求这**6**列随机数据的相关系数矩阵。

命令如下：

```
X=randn(10000,6);
```

```
M=mean(X)
```

```
D=std(X)
```

```
R=corrcoef(X)
```

6.1.6 排序

MATLAB中对向量**X**是排序函数是**sort(X)**，函数返回一个对**X**中的元素按升序排列的新向量。

sort函数也可以对矩阵**A**的各列或各行重新排序，其调用格式为：

[Y,I]=sort(A,dim)

其中**dim**指明对**A**的列还是行进行排序。

若**dim=1**，则按列排；

若**dim=2**，则按行排。

Y是排序后的矩阵，而**I**记录**Y**中的元素在**A**中位置。

6.2 离散傅立叶变换

离散傅立叶变换的实现

一维离散傅立叶变换函数，其调用格式与功能为：

- 1、 **fft(X)**：返回向量 X 的离散傅立叶变换。设 X 的长度(即元素个数)为 N ，若 N 为2的幂次，则为以2为基数的快速傅立叶变换，否则为运算速度很慢的非2幂次的算法。对于矩阵 X ，**fft(X)**应用于矩阵的每一列。

2、 **fft(X,N)**: 计算N点离散傅立叶变换。它限定向量的长度为N，若X的长度小于N，则不足部分补上零；若大于N，则删去超出N的那些元素。对于矩阵X，它同样应用于矩阵的每一列，只是限定了向量的长度为N。

3、 **fft(X,[],dim)**或**fft(X,N,dim)**: 这是对于矩阵而言的函数调用格式，前者的功能与**FFT(X)**基本相同，而后者则与**FFT(X,N)**基本相同。只是当参数**dim=1**时，该函数作用于X的每一列；当**dim=2**时，则作用于X的每一行。

值得一提的是，当已知给出的样本数 N_0 不是2的幂次时，可以取一个 N 使它大于 N_0 且是2的幂次，然后利用函数格式`fft(X,N)`或`fft(X,N,dim)`便可进行快速傅立叶变换。这样，计算速度将大大加快。

相应地，一维离散傅立叶逆变换函数是`ifft`。`ifft(F)`返回 F 的一维离散傅立叶逆变换；`ifft(F,N)`为 N 点逆变换；`ifft(F,[],dim)`或`ifft(F,N,dim)`则由 N 或 dim 确定逆变换的点数或操作方向。

例6-16 给定数学函数

$$x(t)=12\sin(2\pi\times 10t+\pi/6)+6\cos(2\pi\times 60t)$$

取N=128，试对t从0~1秒采样，用fft作快速傅立叶变换，绘制相应的振幅-频率图。

程序如下：

```
N=128;                % 采样点数
T=1;                  % 采样时间终点
t=linspace(0,T,N);    % 给出N个采样时间ti(I=1:N)
x=12*sin(2*pi*10*t+pi/6)+6*cos(2*pi*60*t); % 求各采样点
                        % 样本值x
dt=t(2)-t(1);         % 采样周期
f=1/dt;               % 采样频率(Hz)
X=fft(x);             % 计算x的快速傅立叶变换X
F=X(1:N/2+1);         % F(k)=X(k)(k=1:N/2+1)
f=f*(0:N/2)/N;        % 使频率轴f从零开始
plot(f,abs(F),'-*')   % 绘制振幅-频率图
xlabel('Frequency');
ylabel('|F(k)|')
```

MATLAB统计分析举例

- 统计分析函数
 - `prod(x)`—计算数组**x**的连乘值
 - **`y=sort(x)`**—排序函数
 - 举例1:
 - `rains=[126.8 148.5 173.0 148.4 194.7
208.9;328.8 300.7 268.3 210.5 278.4 321.5;]`
 - `avg_rain=mean(rains)`
 - `avg_rain=mean(avg_rain)`
 - `max_rain=max(rains)`
 - `[max_rain,x]=max(rains)`

MATLAB统计分析

- 统计分析函数
 - `min_rain=min(rains)`
 - `s_sort=sort(rains)`
 - 举例2:
 - `x=[1 2 3 4 5];`
 - `prod(x)`
 - `cumsum(x)`
 - `cumprod(x)`

MATLAB统计分析

- 标准差和协方差
 - 标准差的概念
 - 计算函数标准差函数std
 - rains=[126.8 148.5 173.0 148.4 194.7 208.9
328.8 300.7 268.3 210.5 278.4 321.5]
 - s_dev=std(rains)
 - 协方差的概念
 - 函数cov(x)计算向量x的方差；对于矩阵，
cov(x)对应于协方差矩阵
 - cov(rains)
 - sqrt(cov(rains))

MATLAB统计分析

- 标准差和协方差

- 当x和y是等长度的观测值时，可用`cov(x,y)`计算其协方差，运算等同于`cov([x(:) y(:)])`

- 举例：

- `x=[126.8 148.5 173.0 148.4 194.7 208.9
328.8 300.7 268.3 210.5 278.4 321.5];`

- `y=[120.8 138.5 158.0 148.4 194.7 208.9
328.8 300.7 268.3 210.5 378.4 221.5];`

- `cov(x,y)`

MATLAB统计分析

- 综合分析函数datastats()
 - 函数`ds=datastats(x)`对数据`x`做统计运算，`x`应该是一个实数列向量，如果以复数代入，其虚数部分将被忽略，`ds`的内容如下：
 - `ds.num`—样本容量
 - `ds.max`—最大值
 - `ds.min`—最小值
 - `ds.mean`—平均值
 - `ds.median`—中值
 - `ds.range`—最大值-最小值
 - `ds.std`—标准差

MATLAB统计分析

- 综合分析函数datastats()
 - 函数[xds,yds]=datastats(x,y)对数据x和y同时做统计运算，结果赋给变量xds和yds
 - 举例：
 - x1=[15 14 11 10 9 7 4 3 4 2];
 - x2=[13 17 15 11 13 17 13 11 11 11];
 - da1=datastats(x1')
 - [da1,da2]=datastats(x1',x2')