由数模加油站搜集整理，获取更多数学建模相关资料关注【公众号：数模加油站】

数据预处理——缺失值处理

若对数据求个均值（mean），如果数据中含有缺失值，MATLAB 会返回一个空值（NaN），此时即表示数据中存在缺失值。

MATLAB 提供了以下函数帮助处理缺失值：

(1) 寻找缺失值（ismissing / isnan）

TF = ismissing(A)

A 为输入数组，可以是向量、矩阵或多维数组等。

TF 为输出的逻辑数组，指示数组或表中的哪些元素包含缺失值。TF 的大小与 A 的大小相同。 ‘1’代表缺失值，‘0’则不是。

isnan 用法与 ismissing 相同。

```
% ismissing 寻找缺失值示例
    % 随机生成 6×4 数据矩阵
    A = randi(10,6,4);
    % 替换一些值为缺失值
    A(3:4,2) = nan;
    % 输出数据矩阵
    disp(A)
    % 寻找缺失值，1 代表缺失值所在的位置
    ismissing(A)
```

公众号：数模加油站
QQ群：544457657

对于缺失值处理，我们需要分情况讨论。

如果某个变量或某个样本缺失了 70%以上的数据，那么此时对数据进行填补的话会引入更多的噪声，反而会降低模型的性能，故此时一般直接将该变量或样本删除；

如果缺失的不多，我们可以考虑对缺失值进行填补。

以下重点介绍 MATLAB 中删除和填补缺失值的两个函数：

(2) 删除缺失值（rmmissing）

由数模加油站搜集整理，获取更多数学建模相关资料关注【公众号：数模加油站】

由数模加油站搜集整理，获取更多数学建模相关资料关注【公众号：数模加油站】

[R, TF] = rmmissing(A, dim)

A 为输入数组，可以是向量、矩阵或多维数组等。

dim: 运算维度，默认为 1，删除缺失值所在行，设为 2 则删除缺失值所在列。

R 为删除缺失值条目后的数据。

TF 为已删除条目的指示符。

rmmissing 这个函数比较简单粗暴，只要有缺失值就把缺失值所在行删了，谨慎使用。

(3) 填补缺失值（fillmissing）

方法 1: 常数填充，如用常数 1 填充

F = fillmissing(A, 'constant', 1);

方法 2: F = fillmissing(A,method)

method 是一个字符型参数，指填充缺失值的方法，MATLAB 提供了以下几种方法：

方法（method）	说明
'previous'	上一个非空值
'next'	下一个非空值
'nearest'	最邻近的非空值
'linear'	线性插值
'spline'	三次样条插值
'pchip'	保形三次样条插值

公众号：数模加油站
QQ群：544457657

方法 3: F = fillmissing(A,movmethod,window)

移动窗口填充法，该方法的思想是在缺失值前后开一个“窗口”，用“窗口”内的数据的均值或中位数进行填充

方法（movmethod）说明

'movmean' 窗口长度为 window 的移动均值

'movmedian' 窗口长度为 window 的移动中位数

举个栗子：

x = [4,2,5,3,nan,8,1,5,9]

由数模加油站搜集整理，获取更多数学建模相关资料关注【公众号：数模加油站】

由数模加油站搜集整理，获取更多数学建模相关资料关注【公众号：数模加油站】

```
fillmissing(x,'movmean',7)
```

window 表示窗口长度，必须是一个正整数。 如果 window 是正整数标量，则窗口以当前元素为中心并且包含 window-1 个相邻元素。 如果 window 是偶数，则窗口以当前元素和上一个元素为中心。

插补可方法	方法描述
均值/中位数/众数插补	根据属性值的类型，用该属性取值的平均数/中位数/众数进行插补。
使用固定值	将缺失的属性值用一个常量替换。如广州一个工厂普通外来务工人员的“基本工资”属性的空缺值可以用 2015 年广州市普通外来务工人员工资标准 1895 元/月，该方法就是使用固定值。
最近临插补	在记录中找到与缺失样本最接近的样本的该属性值插补
回归方法	对带有缺失值的变量，根据已有数据和与其有关的其他变量（因变量）的数据建立拟合模型来预测缺失的属性值。
插值法	插值法是利用已知点建立合适的插值函数 $f(x)$，未知值由对应点 x_i 求出的函数值 $f(x_i)$ 近似代替。

公众号：数模加油站 QQ群：544957657 www.mo...

其中，插值法有 Hermite 插值、分段插值、样条插值法，而最主要的有拉格朗日插值法和牛顿插值法。

由数模加油站搜集整理，获取更多数学建模相关资料关注【公众号：数模加油站】