程序设计练习1：线性回归

机器学习

# 介绍

在本练习中，您将实现线性回归，并了解它对数据的作用。在开始这个编程练习之前，我们强烈建议您观看视频讲座并完成相关主题的复习问题。

要开始练习，您需要下载起始代码并将其内容解压缩到您希望完成练习的目录中。如果需要，在开始本练习之前，请使用octave/matlab中的cd命令更改到此目录。

您也可以在课程网站的“环境设置说明”中找到安装octave/matlab的说明。

## 本练习中包含的文件

ex1.m-八度/matlab脚本，引导您完成练习的后续部分ex1 multi.m-八度/matlab脚本ex1data1.txt-一个变量的线性回归数据集ex1data2.txt-多个变量的线性回归数据集submit.m-将解决方案发送到我们的服务器的提交脚本[]warmupexercise.m-八度/matlab中的简单示例函数[]plotdata.m-显示数据集的函数*？？*

[]compute cost.m-计算线性回归成本的函数*？*

[]GradientDescement.m-运行梯度下降的函数*？*

[†]computecostmulti.m-多变量成本函数

[†]GradientDescentmulti.m-多变量梯度下降

[†]FeatureNormalize.m-用于规范功能的函数

[†]normaleqn.m-用于计算正态方程的函数

*？*指示需要完成的文件

†表示可选练习

在整个练习中，您将使用脚本ex1.m和ex1这些脚本为问题设置了数据集，并调用将要编写的函数。您不需要修改其中任何一个。您只需要按照本分配中的说明修改其他文件中的函数。

对于这个编程练习，只需要完成练习的第一部分，就可以用一个变量实现线性回归。练习的第二部分（可选）包括多变量线性回归。

## 在哪里寻求帮助

本课程中的练习使用八度[1]或matlab，这是一种高级编程语言，非常适合数值计算。如果您没有安装八度音阶或matlab，请参考课程网站“环境设置说明”中的安装说明。

在octave/matlab命令行中，键入帮助，然后键入函数名，显示内置函数的文档。例如，帮助图将显示用于绘制的帮助信息。有关八度函数的更多文档，请参见八度文档页面。matlab文档可以在matlab文档页面中找到。

我们还强烈鼓励使用在线讨论与其他学生讨论练习。但是，不要查看其他人编写的任何源代码，也不要与其他人共享源代码。



# 1             简单的八度/matlab函数

ex1.m的第一部分给出了使用八度/matlab语法和作业提交过程的练习。在warmupexercise.m文件中，您将找到一个八度/matlab函数的轮廓。通过填写以下代码，将其修改为返回5 x 5标识矩阵：

A=眼（5）；

完成后，运行ex1.m（假设您在正确的目录中，在octave/matlab提示下键入“ex1”），您将看到类似以下内容的输出：

ANS＝

对角矩阵

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1 | 0 | 0 | 0 | 0 |
| 0 | 1 | 0 | 0 | 0 |
| 0 | 0 | 1 | 0 | 0 |
| 0 | 0 | 0 | 1 | 0 |
| 0 | 0 | 0 | 0 | 1 |

现在，ex1.m将暂停，直到您按下任何键，然后运行分配下一部分的代码。如果要退出，键入ctrl-c将在程序运行过程中停止该程序。

## 1.1           提交解决方案

完成练习的一部分后，您可以通过在octave/matlab命令行中键入submit来提交要评分的解决方案。提交脚本将提示您输入登录电子邮件和提交令牌，并询问您要提交哪些文件。您可以从网页获取分配的提交令牌。

*现在您应该提交解决方案。*

您可以多次提交解决方案，我们只考虑最高分数。



# 2             单变量线性回归

在本练习的这一部分中，您将使用一个变量实现线性回归，以预测食品卡车的利润。假设你是一家餐厅特许经营的首席执行官，正在考虑在不同的城市开设一家新的餐厅。这个连锁店在各个城市都有卡车，你有城市的利润和人口数据。

您希望使用此数据帮助您选择要扩展到下一个城市的城市。

文件ex1data1.txt包含线性回归问题的数据集。第一列是一个城市的人口，第二列是该城市食品卡车的利润。利润负值表示亏损。

ex1.m脚本已设置为为为您加载此数据。

## 2.1           绘制数据

在开始执行任何任务之前，通过可视化了解数据通常很有用。对于这个数据集，您可以使用散点图来可视化数据，因为它只有两个属性要绘制（利润和人口）。（你在现实生活中会遇到的许多其他问题都是多维的，无法在二维图上绘制出来。）

在ex1.m中，数据集从数据文件加载到变量x和y中：

|  |  |
| --- | --- |
| data=load（&apos;ex1data1.txt&apos;）； | %读取逗号分隔的数据 |
| x=数据（：，1）；y=数据（：，2）；m=长度（y）； | %培训实例数 |

接下来，脚本调用plot data函数来创建数据的散点图。您的工作是完成plotdata.m以绘制绘图；修改文件并填写以下代码：

|  |  |
| --- | --- |
| 绘图（x，y，&apos;rx&apos;，&apos;markersize&apos;，10）； | %绘制数据 |
| 伊拉贝尔（10000美元利润）； | %设置Y轴标签 |
| xlabel（“10000年城市人口”）； | %设置X轴标签 |

现在，当您继续运行ex1.m时，我们的最终结果应该如图1所示，带有相同的红色“x”标记和轴标签。

要了解关于plot命令的更多信息，可以在octave/matlab命令提示下键入helpplot，或者在线搜索绘图文档。（要将标记更改为红色“x”，我们将选项“rx”与plot命令一起使用，即plot（..，[此处为您的选项]……，

“RX”；



图1：训练数据散点图

## 2.2           梯度下降

在本部分中，您将使用梯度下降将线性回归参数θ拟合到我们的数据集。

### 2.2.1          更新公式

线性回归的目标是使成本函数最小化。



其中假设hθ（x）由线性模型给出

*H时*（x）=θtx=θ0+θ1x1

记住，模型的参数是θj值。这些是您将调整以最小化成本j（θ）的值。一种方法是使用批梯度下降算法。在批梯度下降中，每个迭代执行更新

（同时更新所有j的θj）。

随着梯度下降的每一步，你的参数θj会更接近最优值，从而达到最低成本j（θ）。

**实施说明：**我们将每个示例作为一行存储在八度/matlab中的x矩阵中。为了考虑截距项（θ0），我们在x中添加了一个额外的第一列，并将其设置为所有列。这允许我们将θ0简单地视为另一个“特征”。

### 2.2.2          实施

在ex1.m中，我们已经建立了线性回归的数据。在下面的几行中，我们为数据添加另一个维度，以适应θ0截距项。我们还将初始参数初始化为0，学习速率alpha初始化为0.01。

|  |
| --- |
| x=[1（m，1），数据（：，1）]；%在x theta=0（2，1）中添加1列；%初始化拟合参数  迭代次数=1500；alpha=0.01； |

### 2.2.3          计算成本（θ）*J*

在执行梯度下降以学习最小化成本函数j（θ）时，通过计算成本来监控收敛性是很有帮助的。在本节中，您将实现一个计算j（θ）的函数，以便检查渐变下降实现的收敛性。

下一个任务是完成computecost.m文件中的代码，该文件是一个计算j（θ）的函数。在执行此操作时，请记住变量x和y不是标量值，而是其行代表训练集中的示例的矩阵。

完成函数后，ex1.m中的下一步将使用θ初始化为零运行computecost，然后您将看到打印到屏幕上的成本。

预计费用为32.07英镑。

*现在您应该提交解决方案。*

### 2.2.4          梯度下降

接下来，您将在文件gradientdepression.m中实现渐变下降。已经为您编写了循环结构，您只需要在每个迭代中提供θ的更新。

当你编程时，确保你理解你正在尝试优化的内容和正在更新的内容。请记住，成本j（θ）是由向量θ而不是x和y参数化的。也就是说，我们通过更改向量θ的值而不是通过更改x或y来最小化j（θ）的值。如果不确定，请参阅本讲义和视频讲座中的方程式。

验证梯度下降是否正常工作的一个好方法是查看j（θ）的值，并检查它是否随每个步骤而减小。GradientDescence.m的起始代码在每次迭代时调用ComputeCost并打印成本。假设您已经正确地实现了梯度下降和计算成本，那么您的j（θ）值应该永远不会增加，并且应该在算法结束时收敛到一个稳定的值。

完成后，ex1.m将使用最终参数绘制线性拟合。结果应该类似于图2：

θ的最终值也将用于预测35000和70000人的利润。请注意，ex1.m中的以下行使用矩阵乘法而不是显式求和或循环来计算预测。这是一个在八度/matlab中实现代码矢量化的例子。

*现在您应该提交解决方案。*

predict1=[1，3.5]\*theta；predict2=[1，7]\*theta；

## 2.3           调试

在实现梯度下降时，需要记住以下几点：•八度/matlab数组索引从一开始，而不是从零开始。如果将θ0和θ1存储在一个称为theta的向量中，值将是theta（1）和theta（2）。

•如果您在运行时看到许多错误，请检查矩阵操作，以确保正在添加和乘法兼容维度的矩阵。使用SIZE命令打印变量的维度将有助于调试。

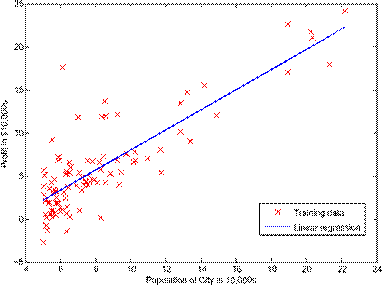


图2：线性回归拟合的训练数据

•默认情况下，octave/matlab将数学运算符解释为矩阵运算符。这是大小不兼容错误的常见来源。如果你不想矩阵乘法，你需要添加“点”符号来指定这个到八度/matlab。例如，A\*B做矩阵乘法，A\*B做元素乘法。

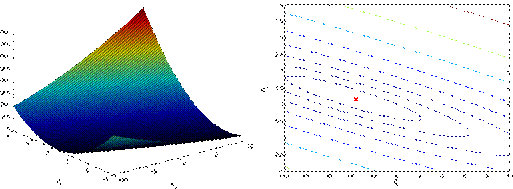
## 2.4           可视化（θ）*J*

为了更好地理解成本函数j（θ），现在您将在θ0和θ1值的二维网格上绘制成本。您将不需要为此部分编写任何新的代码，但您应该了解您编写的代码是如何创建这些图像的。

在ex1.m的下一步中，有一些代码设置为使用您编写的computecost函数在网格值上计算j（θ）。

|  |
| --- |
| %初始化J到0的矩阵的值  Jvals=零（长度（theta0vals），长度（theta1VALS）；  填充Ji=1时的值：j=1时的长度（θ0值）：长度（θ1值）t=[θ0VAL（i）；θ1VARS（J）；  Jvals（i，j）=computecost（x，y，t）；结束  结束 |

执行这些行之后，您将拥有一个由j（θ）值组成的二维数组。然后脚本ex1.m将使用这些值使用surf和contour命令生成j（θ）的曲面和等高线图。图应该类似于图3：



                                 （a）表面（b）轮廓，显示最小值

图3：成本函数j（θ）

这些图的目的是向您展示j（θ）如何随θ0和θ1的变化而变化。成本函数j（θ）为碗形，具有全局最小值。（在等高线图中比在三维曲面图中更容易看到）。这个最小值是θ0和θ1的最佳点，梯度下降的每一步都靠近这个点。



# 可选练习

如果你成功完成了上述材料，祝贺你！现在您了解了线性回归，应该能够在自己的数据集中开始使用它。

对于这个编程练习的其余部分，我们包含了以下可选练习。这些练习将帮助您对材料有更深的了解，如果您能够这样做，我们也鼓励您完成这些练习。



# 3             多元线性回归

在这部分中，您将使用多个变量实现线性回归来预测房屋的价格。假设你在卖房子，你想知道一个好的市场价格是多少。这样做的一个方法是首先收集最近售出的房屋的信息，并制作一个房价模型。

ex1data2.txt文件包含俄勒冈州波特兰市的一套住房价格培训。第一列是房子的大小（以平方英尺为单位），第二列是卧室的数量，第三列是房子的价格。

ex1 multi.m脚本已设置为帮助您逐步完成此练习。

## 3.1           特征归一化

EX1multi.m脚本将从这个数据集中加载和显示一些值开始。通过查看这些值，可以发现房子的大小大约是卧室数量的1000倍。当特征的数量级不同时，首先执行特征缩放可以使梯度下降更快地收敛。

您在这里的任务是完成featurenormalize.m到中的代码

•从数据集中减去每个特征的平均值。

•减去平均值后，再将特征值按其各自的“标准偏差”进行缩放（除）。

标准偏差是一种测量特定特征值范围内有多少变化的方法（大多数数据点将位于平均值的±2标准偏差范围内）；这是取值范围（最大-最小）的替代方法。在Octave/Matlab中，可以使用“std”函数计算标准偏差。例如，在featurenormalize.m中，数量x（：，1）包含训练集中x1（房屋大小）的所有值，因此std（x（：，1））计算房屋大小的标准偏差。在调用featurenormalize.m时，与x0=1对应的1的额外列尚未添加到x中（有关详细信息，请参阅ex1 multi.m）。

您将对所有特性进行此操作，并且您的代码应该与所有大小的数据集（任何数量的特性/示例）一起使用。注意，矩阵X的每一列对应一个特征。

*现在您应该提交解决方案。*

**实施说明：**规范化特征时，重要的是存储用于规范化的值-用于计算的平均值和标准偏差。从模型中学习参数后，我们常常想预测我们以前没有见过的房屋的价格。给定一个新的x值（起居室面积和卧室数量），我们必须首先使用之前从训练集中计算出的平均值和标准偏差将x标准化。

## 3.2           梯度下降

以前，您在单变量回归问题上实现了梯度下降。唯一的区别是矩阵X中还有一个特征，假设函数和批梯度下降更新规则保持不变。

您应该在computecostmulti.m和gradientdescentmulti.m中完成代码，以实现具有多个变量的线性回归的成本函数和梯度下降。如果前一部分（单变量）中的代码已经支持多个变量，那么您也可以在这里使用它。

确保您的代码支持任何数量的特性，并且向量化良好。您可以使用“大小（x，2）”来确定数据集中有多少功能。

*现在您应该提交解决方案。*

|  |
| --- |
| **实施说明：**在多变量情况下，成本函数也可以用以下矢量化形式表示：    哪里                                               （1））-（1）*T*  -（x）  -（x（2））-*T*  *X*=..γγ*.*  γ  γ  -（  当你使用诸如八度/matlab之类的数值计算工具时，矢量化版本是有效的。如果你是矩阵运算的专家，你可以向自己证明这两种形式是等价的。 |

### 3.2.1          可选（未分级）练习：选择学习率

在本部分练习中，您将尝试数据集的不同学习速率，并找到快速收敛的学习速率。您可以通过修改ex1 multi.m和更改设置学习率的代码部分来更改学习率。

ex1的下一阶段multi.m将调用您的梯度下降.m函数，并以所选的学习速率运行大约50次梯度下降。函数还应返回向量j中j（θ）值的历史记录。在上一次迭代之后，ex1 multi.m脚本根据迭代次数绘制j值。

如果您在一个良好的范围内选择了一个学习率，那么您的图类似于图4。如果你的图表看起来非常不同，特别是如果你的j（θ）值增加或甚至爆炸，调整你的学习速度，然后再试一次。我们建议在对数尺度上尝试学习率α的值，以约为前一个值（即0.3、0.1、0.03、0.01等）3倍的乘法步进。如果这有助于您在曲线中看到整体趋势，那么您可能还需要调整正在运行的迭代次数。

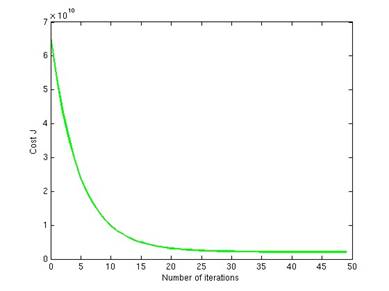


图4：具有适当学习速率的梯度下降的收敛性

**实施说明：**如果你的学习速度太大，j（θ）可能会发散并“爆炸”，导致数值太大，无法进行计算机计算。在这种情况下，octave/matlab将倾向于返回nan。Nan代表“非数字”，通常由涉及−∞和+∞的未定义操作引起。

|  |
| --- |
| **八度/matlab提示：**为了比较不同的学习速率对收敛性的影响，可以在同一个图上绘制几个学习速率的J。在Octave/Matlab中，这可以通过在绘图之间使用“保持”命令多次执行梯度下降来实现。具体来说，如果您尝试了三个不同的alpha值（您可能应该尝试更多的值），并将成本存储在j1、j2和j3中，则可以使用以下命令将其绘制在同一个图上：  绘图（1:50，j1（1:50），“b”）；保持；  绘图（1:50，J2（1:50），“R”）；绘图（1:50，J3（1:50），“K”）；  最后的参数“b”、“r”和“k”为绘图指定不同的颜色。 |

注意随着学习速率的变化收敛曲线的变化。当学习率很小时，你会发现梯度下降需要很长的时间才能收敛到最佳值。反之，学习速度大，梯度下降可能不收敛，甚至可能发散！

使用您找到的最佳学习率运行ex1multi.m脚本运行渐变下降直到收敛，以找到θ的最终值。接下来，使用θ的值来预测1650平方英尺3间卧室的房子的价格。稍后，您将使用值检查法方程的实现。做这个预测时，不要忘记规范化您的特性！

*您不需要为这些可选（未分级）练习提交任何解决方案。*

## 3.3           正态方程

在讲座视频中，您了解到线性回归的闭式解是



使用这个公式不需要任何特征缩放，您将在一个计算中得到一个精确的解决方案：像梯度下降一样，在收敛之前没有“循环”。

在normaleqn.m中完成代码，使用上面的公式计算θ。记住，虽然您不需要扩展特性，但我们仍然需要在x矩阵中添加一列1，以获得截距项（θ0）。ex1.m中的代码将为您添加1到x的列。

*现在您应该提交解决方案。*

*可选（未分级）练习：*现在，一旦你用这个方法找到θ，用它来预测一个1650平方英尺的三居室房子的价格。您应该发现，它给出的预测价格与您使用带有梯度下降的模型拟合获得的值相同（见第3.2.1节）。

# 提交和评分

完成任务的各个部分后，请务必使用提交功能系统将您的解决方案提交给我们的服务器。下面是这个练习的每个部分的评分方式的细目。

|  |  |  |
| --- | --- | --- |
| **部分** | **提交文件** | **点** |
| 热身运动 | 热身运动.m | 10点 |
| 计算一个变量的成本 | 计算器 | 40点 |
| 一个变量的梯度下降 | 梯度下降.m | 50点 |
| 总分 |  | 100点 |

## 可选练习

|  |  |  |
| --- | --- | --- |
| **部分** | **提交文件** | **点** |
| 特征归一化 | 特性normalize.m | 0点 |
| 计算多重成本  变量 | 计算机成本多.m | 0点 |
| 多变量梯度下降 | 渐变色多.m | 0点 |
| 正态方程 | 正态qm | 0点 |

您可以多次提交解决方案，我们只考虑最高分数。

[〔1〕](" \l "_ftnref1" \o ")八度是matlab的免费替代品。对于编程练习，您可以自由使用八度或matlab。