编程练习 1：线性回归

机器学习

# 介绍

在本练习中，您将实现线性回归，并看到它对数据起作用。在开始此编程练习之前，我们强烈重新交流- 修补观看视频讲座并完成相关主题的复习问题。

要对练习进行压缩，you 您需要下载初学者代码并将其内容解压缩到you 您希望完成练习的目录中。如果需要，在开始此练习之前，请使用Octave/MATLAB中的cd 命令更改为此目录。

find instructions您还可以在课程网站的 the can"en- virave/MATLA B"中找到安装 installing八度/MATLAB的说明。

## 本练习中包含的文件

ex1.m - Octave/MATLAB脚本，用于you through the exercise 练习后期部分的multi.m - Octave/MATLAB Octave/MATLAB 脚本，用于练习的后期部分 ex1data1.txt - 具有一个变量ex1data2.txt的线性回归数据集 - 具有多个变量提交.m 的线性回归数据集-将您的解决方案发送到我们的服务 rs 的提交脚本rs [*] [ ]*暖上练习.m -八度内的简单示例函数 \**• •*绘图Data.m -显示数据集的函数

[*\**] 计算成本.m - 计算线性回归成本的函数

[*\**] 梯度下降.m - 运行梯度下降的函数

*[ ]*]computeCostMulti.m - multiple v可阿里的成本函数

*[ ]*]格拉迪ntSssMulti.m - 格拉迪nt dessnt 为 multiple v可安迪斯

*[ ]*]feature规范化.m - Function 使 normalize feaa tures

*[ ]*]规范lEqn.m - Function 计算 compute the 法线方程

*•*指示您需要完成的文件

*•*指示可选练习

在整个练习中，您将使用脚本 ex1.m 和 ex1多m。

这些脚本为 the make问题 set设置数据集，并 calls调用 functions that 要编写的函数。您无需修改其中任何一个。You 您只需要by 按照此分配中的说明修改其他文件中的函数。

对于此编程练习，您 are only只需要 to完成 the练习的第一部分，即可使用一个变量实现线性回归。练习的第二部分是可选的，它涵盖了具有多个变量的线性回归。

## 在哪里获得帮助

本课程中的练习使用八角形 1[1](#_bookmark0)或 MATLAB，这是一种高级程序明语言，非常适合数值计算。如果您you 没有have 安装八度或MATLAB，请参阅课程网站"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func-

tion名称显示内置函数的文档。例如，帮助绘图将显示用于绘图的帮助信息。八度扫描函数的Further 文档可以在found at the [八度文档页](http://www.gnu.org/software/octave/doc/interpreter/)中找到。MAT-LAB 文档可在be [MATLAB 文档页](http://www.mathworks.com/help/matlab/?refresh=true)中找到。

我们还强烈建议 encourage使用 the在线**讨论** to with与其他学生讨论前讨论。但是， do不要查看 at其他人编写by 的任何源代码 code或与他人共享您的源代码。

# 简单的八度/MATLAB功能

ex1.mgives 的第一部分为您提供了使用八度/MATLAB 语法和家庭作业提交过程的实践。在文件暖上练习.m中，您将找到八度/MATLAB函数的轮廓。 通过填写以下代码修改它以返回 5 x 5 标识矩阵by filling in the following：

A = 眼睛（5）;

1八度是MATLAB的免费替代品。对于编程练习，您是自由的

使用八度或 MATLAB。

完成后，运行ex1.m（假设您处于正确的 di-recre，在八角形/MATLAB 提示符处键入"ex1"），您应该会看到类似于以下内容的输出： ex1

岁 =

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，直到您按任意键，然后运行分配的下一部分的代码。如果要you 退出，键入 ctrl-c 将在运行中停止程序。

## 提交解决方案

完成部分练习后，您可以通过submit 在 Octave/MATLAB 命令行键入提交来提交分级解决方案。子任务脚本将提示您输入您的登录电子邮件和提交令牌，并询问您要 submit 的文件。可以从网页上获取分配的提交令牌。

*您现在应该提交您的解决方案。*

您可以多次提交您的解决方案，我们将will只考虑最高分。

# 一个变量的线性回归

在本练习的这一部分中you ，您将使用一个变量实现线性回归，以预测食品卡车的利润。假设您you 是一家餐厅特许经营的首席执行官，并正在考虑不同的城市开设一家分店。该连锁店已经在各个城市拥有卡车，并且您拥有来自城市的利润和人口数据。

您希望使用此数据来帮助选择要扩展到下一个城市。

文件ex1data1.txt 包含我们的线性回归概率的数据集-

莱姆.第一列是城市的人口，第二列是城市中食品卡车的利润。利润的负值表示亏损。

ex1.m 脚本已设置为为您加载此数据。

## 绘制数据

在开始执行任何任务之前，通过可视化来了解数据通常很有用by 。对于此数据集，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（利润和总体）。（许多其他问题，在现实生活中，你会是多维的，不能be 在二维图上绘制。

在ex1.m中，数据集从数据文件加载到变量 *X*

和*y*：

培训示例数百分比

读取逗号分隔数据的百分比

数据 = 负载（"ex1data1.txt"）;);

X = 数据（：，1）;y = 数据（：2）;m = 长度（y）;

接下来，脚本调用plotData 函数以创建数据的散点图。你的工作是完成绘图Data.m来绘制绘图;修改文件并填写以下代码：

百分比设置y*+*轴标签

xlabel("在万城市人口"); 百分比设置 x*+*轴标签

打印数据百分比

[ （x，y， 'rx'， ''Rx','标记大小'',10);ylabel("利润在 10，000 美元"）;);

现在，当您继续运行ex1.m时，我们的最终结果应如图 [1 所示，](#_bookmark1)具有相同的红色"x"标记和轴标签。

要了解有关绘图命令的更多详细信息，you 可以在help plot at the

八度/MATLAB命令提示或 to search在线 for搜索绘图文档。. （为了将标记更改为红色"x"，我们使用选项"rx"和绘图命令，即绘图（.，[此处的选项options ],..,'rx'）;rx’); )

25

20

15

10

利润在 10，000 美元

5

0

•5

4 6 8 10 12 14 16 18 20 22 24

10，000年代的城市人口

图 1：训练数据的散射图

## 梯度下降

在本部分中，您将使用梯度下降将线性回归参数 *\** 拟合到我们的数据集中。

### 更新方程

线性回归的目标是最小化成本函数

*J* (*θ*) = 1 区（*H*(*x*(*i*))*− y*(*i*))2

*i*±1

*米*

2*米*

*Ⅰ*

其中假设*h=*（*x*）由线性模型给出

*h-*  *（x）* - *"T* *x"* '0' *θ* ' 1*x*1' *θ*

回想一下，模型的参数是*μ jj* 值。这些是values you 您将调整的值，以尽量减少成本 *J*  （*\**）。执行此操作的一种方法是使用批处理梯度下降算法。在批处理梯度下降中，每次迭代执行更新

1 区 (*i*) (*i*) (*i*)

*米*

*Ⅰ*:=*θ − α* (*H*(*Ⅹ*) *− y* )*x* （同时更新*Ⅰ*对于所有 *j*)*.*

*i*±1

*J*

*J*

*米*

*Ⅰ*

*J*

*J*

closer to随着梯度下降的每一步 of，您的参数 *=j* 更接近实现最低成本 *J*  （*μ*）的最优值。

**实现说明：**我们将每个示例存储在 Octave/MATLAB 中的X 矩阵中作为行in the 。要考虑截取术语 （*=*0），我们将第一列添加到 X 并将其设置为所有列。 这使我们能够将 *±*0 视为 simply另一个"功能"。

### 实现

在ex1.m中，我们已经设置了线性回归的数据。在以下行中，我们将另一个维度添加到数据中，以适应 *±*0 截取术语。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（米，1），数据（：，1）\*;百分比 将一列的列添加到 x ta = 零（2，1）;初始化拟合参数的百分比

迭代 = 1500;

α = 0.01;

* + 1. **计算成本***J*  （*\**）

当您执行梯度下降以学习最小化成本函数*J*  （*+*）， 通过计算成本来监视收敛是很有帮助by 的。在本节中，您将 will实现一个函数来计算*J* （*\**），以便您可以 you can检查梯度下降实现的收敛性。

下一个任务是完成文件计算Cost.m中的代码，这是计算J（\*）*θ*的函数*J* 。you 执行此操作时，请记住变量 *X* 和 *y* 不是标量值，而是其行表示训练集中的示例的矩阵。

完成you have 函数后，ex1.m 中的下一步将使用*+*初始化为零后运行计算成本，您将看到打印到 screen的成本。

你应该看到32.07的成本。

*您现在应该提交您的解决方案。*

### 梯度下降

接下来，您将在文件渐变下拉下实现渐变下降。 循环结构已为您编写，您只需在每个迭代中 updates提供 only need to supply the  *\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_*

you 编程时，请确保您you 了解您尝试选择的内容- mize 和正在更新的内容。请记住，成本 *J* （\**θ*） 是参数- 由向量 *\** terter， , 而不是 *X* 和 *y*。也就是说，我们通过更改向量*\** 的值而不是by 通过更改*X*或 *y*来最小化 *J* （*μ）* 的值。 if如果您 are不确定，请参阅本讲义中的方程数和 to the视频讲座。

验证梯度下降是否正常工作的一个好方法是查看 *J*  （*+*）的值，并检查其是否随着每个步骤而减小。梯度 Descent.m的起始代码在每次迭代时调用计算成本并打印成本。假设您have 正确实现了梯度下降和计算成本，则  *J* （*+*） of the value的值 should不应增加， 并且应在 a算法结束时收敛到稳定值.

you 完成后，ex1.m 将使用最终参数绘制线性拟合。结果应类似于图 [2：](#_bookmark2)

\**的最终值*也将用于预测

地区有35，000人和70，000人。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）ing来计算预测的方式。这是八度/MATLAB 中代码矢量化的示例。

*您现在应该提交您的解决方案。*

预测1 = [1， 3.5] = 这;

预测2 = [1， 7] = 这;

## 调试

实现梯度下降时，需要记住一些事项：

* 八度/MATLAB数组索引从 1 开始，而不是零。如果在称为theta的.矢量中 stor-ing ing  *= 0*0 和 *^*1，则值将是be ta（1） 和 ta（2）。1
* 如果在 are运行时看到许多错误 at，请检查 your矩阵操作，以确保添加和乘以可匹配的ible维度矩阵。使用大小命令打印变量的维度size 将帮助您you进行调试。

25

训练数据 线性回归

20

15

10

利润在 10，000 美元

5

0

•5

4 6 8 10 12 14 16 18 20 22 24

10，000年代的城市人口

图 2：具有线性回归拟合的训练数据

* 默认情况下，八度/MATLAB 将数学运算符解释为be 矩阵运算符。 This这是 a大小不兼容错误的常见来源 of。如果 you 不需要want 矩阵乘法，则需要添加"点"表示法以将其指定为 Octave/MATLAB。例如example, ，A\_B执行矩阵乘法，而 A.\_B 执行元素乘法。
  1. **可视化***J* （*\**）

为了更好地理解成本函数*J*  （*+*）， 现在您将now在*^*0和 *±*1值的二维网格上绘制成本。您无需为此部分编写任何新代码，但应了解you 编写的代码是如何创建这些映像的。

在ex1.m的下一步中，使用您编写的计算成本乐趣的拐点设置代码，以在值using the 网格上计算*J*  （*\**） 。

将 J 华尔兹初始化到 0 的矩阵的百分比

J华尔兹 = 零（长度（ta0 华尔兹）长度（ta1 华尔兹）;长度（ta1华尔兹）;

百分比 填写 J瓦尔斯

对于 i = 1：长度（ta0 瓦尔斯）j) = 1：长度（ta1 华尔兹）)

t - [ta0]华尔兹（i）; ta1 华尔兹（d）\*;J 华尔兹（i.j）- 计算成本(（x、y、t）;

结束

结束

执行这些行后，您将具有一个*J*  （*+*） 值的二维数组。然后，脚本ex1.m 将让我们e 这些值，使用冲浪和轮廓命令生成*J*  （*+*）的表面和轮廓图。绘图应类似于图 [3：](#_bookmark3)

|  |  |
| --- | --- |
|  | 4 |
| 3.5 |
| 800  700 | 3 |
| 600 | 2.5 |
| 500 | 2 |
| 400 | 1 |

300

200

100

0

4

3

2 5

1 0

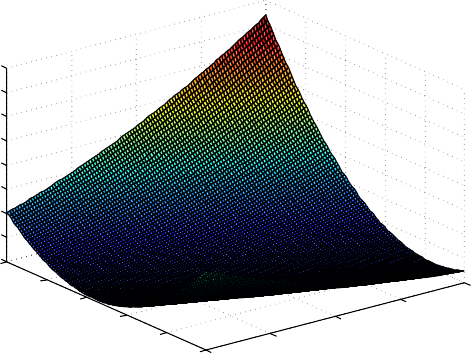
0 ×5

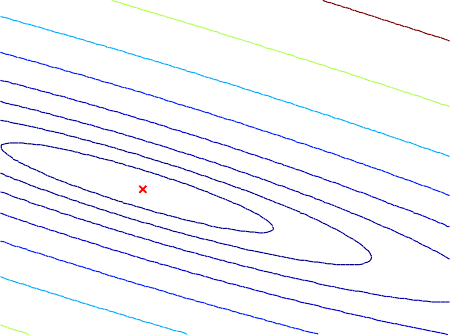
−1−10

•

1 =0

1.5

1



•

0.5

0

10

•0.5

•1

#10 +8 +6 +4 +2 0 2 4 6 8 10

•0

（a）表面 （b） 轮廓，显示最小

图3：成本函数 *J*  （*\**）

这些图形的目的是向您展示that *J*  （*\**） 如何随 \*0 和*\**1中的更改*θ*而变化。成本函数 *J*  （*+*）是碗形的，具有全局微型数字。（这比在 3D 曲面图中更容易在轮廓图中看到）。此最小值是 *±*0 和 *±*1的最佳点，梯度下降的每个步骤都更接近此点。

# 可选练习

如果您已成功完成上述材料，恭喜您！现在，您了解线性回归，应该能够开始在自己的数据集上使用它。

对于本编程练习 the的其余部分 of，我们 have包括以下可选练习。这些练习将帮助您 help you gain a更深入地了解材料，如果您you 能够这样做，我们鼓励encourage 您完成它们。

# 具有多个变量的线性回归

在本部分中，您将实现具有多个变量的线性回归，以预测房价。 假设you 你卖你的房子，你想知道want 什么是良好的市场价格。这样做的一个方法是首先收集有关最近房屋的信息，并建立房价模型。

文件ex1data2.txt 包含一套培训型的端口房价-

土地，俄勒冈州。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，以帮助您逐步完成此

运动。

## 功能规范化

ex1 multi.m 脚本将从by 加载和显示此数据集中的一些值开始。通过查看这些值，请注意，房屋大小大约是卧室数量的 1000 倍 the number of。当要素 differ因 of磁石的阶而异时，每个成形特征缩放首先可以使梯度下降收敛much 更快。

您在这里的任务是完成功能Normalize.m to到

* 从 the数据集中减去每个要素 value的平均值。
* 减去平均值后，额外缩放（除以）要素值，按其各自的"标准差"。

的 标准 偏差 是 a 方式 的 测量 如何 多 变化 有 是 在 范围 的 值 的 a 特定 特征 （大多数 数据 点 将 谎言 在 2 标准 偏差 的 均值）; 这 是 a 替代 自 采取 范围 的 值 （最高分钟）。 在 八度/MATLAB， 你 可以 使用 “小时” 功能 自 计算 标准 偏差。 对于 例子 里面 功能规范化.m, 数量 X（：，1） 包含 所有 值 的 *x*1 （房子） 大小） 在 培训 设置 所以 （X（：，1）） 计算 标准 偏差 的 房子 大小。 在 时间 功能规范化.m 是 叫 额外 列 的 1 的 相应 自 *x*0 = 1 Hsa 不 但 被 添加 自 X （参见 前1 多米 对于

*±*

详细信息）。

您将对所有要素执行此操作，并且代码应使用各种大小的数据集（任意数量的要素/示例）。请注意，矩阵 X的每个列对应于一个要素。

*您现在应该提交您的解决方案。*

**实现说明：**规范化要素时，存储以存储用于规范化的值非常重要 -*平均值价值*和用于计算的*节-达德偏差*。 从模型中了解参数后，价格的我们有我们经常经常想要预测我们以前从未见过的房价。 给定一个新的**Ⅹ**值（客厅面积和卧室数量），我们必须首先使用均值和 st和ard 偏差使**x**规范化

我们以前从训练集中计算过。

## 梯度下降

以前，您对单变量回归问题实现了梯度下降。现在唯一的区别是矩阵 X 中还有一个要素X。下丘同色函数和批处理梯度下降更新规则保持不变。

implement the您应该在计算成本多和 and梯度下线多m中完成 the代码，以实现线性回归 with的成本函数和梯度下降 for

多个变量。如果 your前一部分（单个变量）中的代码已支持多个变量，您也可以在此处使用它。

确保您 your的代码支持 any任意数量的 of功能，并且具有良好的矢量化。您可以使用 'size（X， 2）'来了解how 数据集中有多少要素。

*您现在应该提交您的解决方案。*

使用数字时，矢量化版本是有效的

计算工具，如八度/MATLAB。 如果您是矩阵操作专家，则可以向自己证明这两种形式是等效的。

*.*

·

·

·

*和*(1)

*和*(2)

.

*y*（*米*）

·

·

*和y* =

·

·

·

[*x*(1)）*T* |

[*x*(2)）*T* |

.

[*x*（*m*）T*T* =

·

*Ⅹ*|

·

其中

*T*

1

*J*(*s*） *y* *y* （(*x x*  *•* *y*）（ x x 和 y ）(*xs* *•* *和*） 2*m*

**实施说明：**在多变量的情况下，成本函数可以

也以以下矢量化形式编写：

### 可选（未分级）练习：选择学习率

在练习的这一部分中you ，您将尝试数据集的不同学习速率，并找到快速融合的学习速率。You 您可以通过修改 ex1 multi.m 并更改设置学习速率的代码部分来更改学习速率。

ex1 multi.m的下一阶段将调用您的梯度流.m func- tion，并在所选学习速率下运行大约 50 次迭代的梯度下降。T他函数还应返回向量中 *J*  （*+*） 值的历史记录

J. .上次迭代后，ex1 multi.m脚本根据迭代次数绘制 J 值。

如果在良好范围内选择了学习速率，则绘图看起来类似于图 [4。](#_bookmark4)如果你的图形看起来非常不同，特别是如果你的value of *J*值增加*θ*，甚至爆炸，调整你的学习速率，再试一次。我们以ommend大约 3 倍的倍数步数value （即 0.3、0.1、0.03、0.01 等）在日志刻度上尝试学习 rate *=* 的值。you 如果这将有助于您在曲线中看到总体趋势，also want 则可能需要may 调整正在运行的迭代次数。

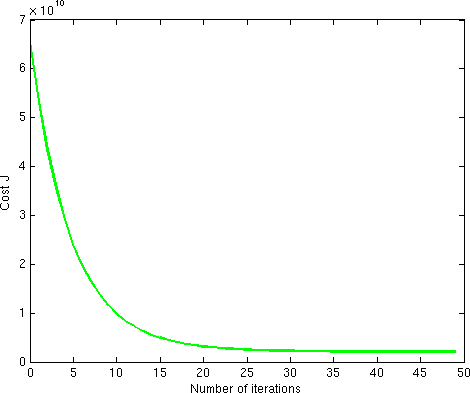


图 4：梯度下降的收敛与适当的学习速率

最后参数"b"，"r"和"k"为

情节。

**八度/MATLAB提示：**比较比较不同学习学习方式

速率影响收敛，在同一图上绘制几个学习速率的 J 是很有帮助的。 在八度/MATLAB 中，可以通过是在绘图之间多次由使用"保持"命令执行梯度下降来实现。具体，如果您你已经尝试了三个不同的 阿 尔 法 使用自值（您可能应该尝试大于值此值的值），并将成本存储在J1×J2 J2和J3中，则可以使用可以以下以下命令绘制相同的上相同数字：

\*（1：50，J1（1：50），"b"）;

坚持;

情节（1：50，J2（1：50），'r'）;

情节（1：50，J3（1：50），"k"）;

**实施说明：**如果您的你学习速率是太大J  (*+*） 可能会偏离边缘和"爆炸"，导致值太大，无法计算机计算。 在这些情况下，八度/MATLAB 将倾向于返回 南。南代表"不是数字"，通常由涉及*\** 和\**∞*的未定义的操作引起。

注意随着 in the学习速率的变化而出现收敛曲线 as the的变化。学习速度小时，您应该发现梯度下降需要很长时间才能收敛到最佳值。相反，随着学习速率大，梯度下降可能不会收敛，甚至可能偏离！

使用您找到的最佳学习速率，运行ex1 multi.m 脚本

运行梯度下降，直到收敛，以找到*|*的最终值。接下来，使用此值value of *来*预测 1650 平方英尺和 3 贝德罗欧姆的房子的价格。稍后将使用值来检查正常方程的实现。做出此预测时，不要忘记规范化功能you make this！

*您无需为这些可选（未分级）练习提交任何解决方案。*

## 法线方程

在讲座视频中，您了解到线性回归的闭合形式解决方案是

*θ* = *XT X −*1 *XT yy.*

（ ) ( )

is no使用此公式不需要任何特征缩放，您将在一个计算 there中获得will get an确切的解决方案：没有"循环直到收敛"，如梯度下降。

在普通 Eqn.m中完成代码，将上述公式用于 calcu-

晚*=*。请记住，虽然您不需要缩放要素，但我们仍然需要向 X 矩阵添加 1 的列，以便具有拦截 trm （=*θ*0）。 ex1.m中的代码将为您添加 1 的列到 X。

*您现在应该提交您的解决方案。*

*可选（未分级）练习：*现在，一旦您have 发现*使用*此方法，使用它对一个 1650 平方英尺的带 3 间卧室的房子进行价格预测。You 您应该发现，给出的预测价格与使用适合 with梯度下降 the的模型获得的值相同（在第 [3.2.1](#_bookmark4) 节中）。

# 提交和分级

完成分配的各个部分后，请务必使用提交功能系统将您的解决方案提交到我们的服务器。以下是本练习的每个部分的评分的细目。

|  |  |  |
| --- | --- | --- |
| **部分** | **已提交文件** | **点** |
| 热身锻炼 | 暖上练习.m | 10 分 |
| 计算一个变量的成本  一个变量的渐变下降 | 计算成本.m  梯度下降.m | 40分  50分 |
| 总积分 |  | 100 点 |

### 可选练习

|  |  |  |
| --- | --- | --- |
| **部分** | **已提交文件** | **点** |
| 多个变量 的功能 规范化计算 成本  多重渐变下降  变量 | 功能规范化.m  计算成本多亿梯度下降多米 | 0点  0点  0点 |
| 法线方程 | 正常Eqn.m | 0 点 |

您可以多次提交您的解决方案，我们将只考虑最高分。