作者	**生姜 DrGinger**
脚本	**生姜 DrGinger**
视频	**崔崔 CuiCui**
开源学习资源	https://github.com/Visualize-ML
平台	https://www.youtube.com/@DrGinger_Jiang https://space.bilibili.com/3546865719052873 https://space.bilibili.com/513194466

12.2 椭圆视角看主成分分析

本节你将掌握的核心技能：

- 用椭圆或椭球来几何直观地描述数据分布。
- 对每列数据去均值，使质心移动到原点。
- 协方差矩阵描述数据结构。
- 谱分解提取数据的主轴方向及方差。
- 方阵的迹是主对角线上所有元素的总和。
- 将数据投影到特征向量方向，实现信息压缩。
- 通过旋转坐标轴 (主成分方向) “摆正”数据形状。

把数据矩阵想象成一朵云，而描述这朵云的合适的几何形状可以是椭圆、椭球。主成分分析就是找到合适的坐标系把代表数据的椭圆、椭球摆正；然后在新的坐标系中完成投影。

⚠ 注意，使用主成分分析时，我们假设数据服从多元高斯分布。“数学不难”系列的《概率统计不难》将会专门讲解多元高斯分布。

我们能不能找到一个新的坐标轴，使得数据变得“摆正”？并且“压平”在最有用的方向上。这就是主成分分析的几何直觉。

我们用数据矩阵计算协方差矩阵；用协方差矩阵描述椭圆、椭球。用特征值分解 (谱分解) 找到椭圆、椭球的主轴 (主元方向)。用正交投影将数据降维到主元方向。从这个思路来看，我们将会用到本书前文介绍的谱分解、正交投影、格拉姆矩阵等工具。

主成分分析步骤

主成分分析的步骤如下：

- 数据矩阵列中心化 (标准化，如果不同特征上方差差异过大)；

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。
代码及 PDF 文件下载：https://github.com/Visualize-ML
本书配套微课视频均发布在 B 站——生姜 DrGinger：https://space.bilibili.com/513194466
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

- 计算协方差矩阵；
- 对协方差矩阵特征值分解；
- 按特征值从大到小顺序排列对应特征向量；
- 选出主元方向 (最大特征值对应的向量)；
- 将数据矩阵投影到主元方向。

⚠ 注意，计算协方差矩阵时，数据中心化 (多数情况下) 已经包含在内。

下面让我们逐步来看主成分分析。

数据矩阵

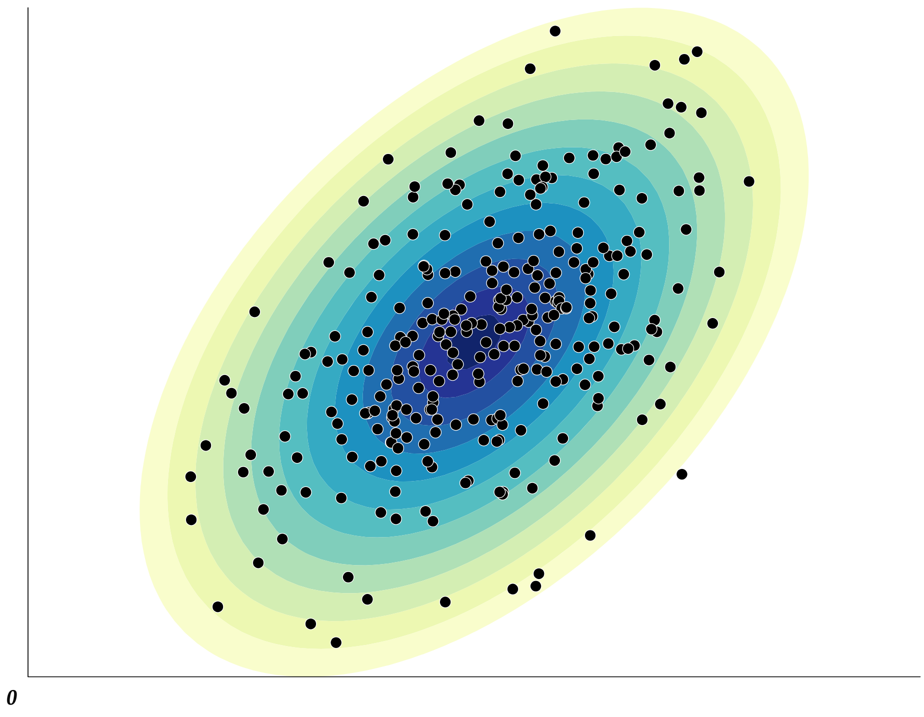
假设我们有一组二维数据点，对应的数据矩阵为

$$\boldsymbol{X}_{300\times 2} = \begin{bmatrix} 2.272 & 1.672 \\ 2.031 & 3.272 \\ \vdots & \vdots \\ 4.894 & 3.716 \end{bmatrix} \tag{1}$$

数据矩阵有 300 行，即 300 个样本点。

如图 1 所示，数据散点像一团斜着的椭圆云。

图 1 中椭圆实际上是马氏距离等高线，马氏距离是本书第 13 章、第 4 节要介绍的话题。



本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。
代码及 PDF 文件下载：https://github.com/Visualize-ML
本书配套微课视频均发布在 B 站——生姜 DrGinger：https://space.bilibili.com/513194466
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

图 1. 旋转椭圆“数据云”

如图 2 所示，数据的质心就是旋转椭圆的中心，通过如下矩阵乘法，我们计算得到了质心具体位置

$$\boldsymbol{\mu}=\begin{bmatrix}\mu_1\\\mu_2\end{bmatrix}=\frac{\boldsymbol{X}^{\mathrm{T}}\boldsymbol{1}}{n}=\frac{1}{300}\times\begin{bmatrix}2.272 & 1.672\\2.031 & 3.272\\\vdots & \vdots\\4.894 & 3.716\end{bmatrix}^{\mathrm{T}}@\begin{bmatrix}1\\1\\1\\1\end{bmatrix}=\begin{bmatrix}4\\3\end{bmatrix} \tag{2}$$

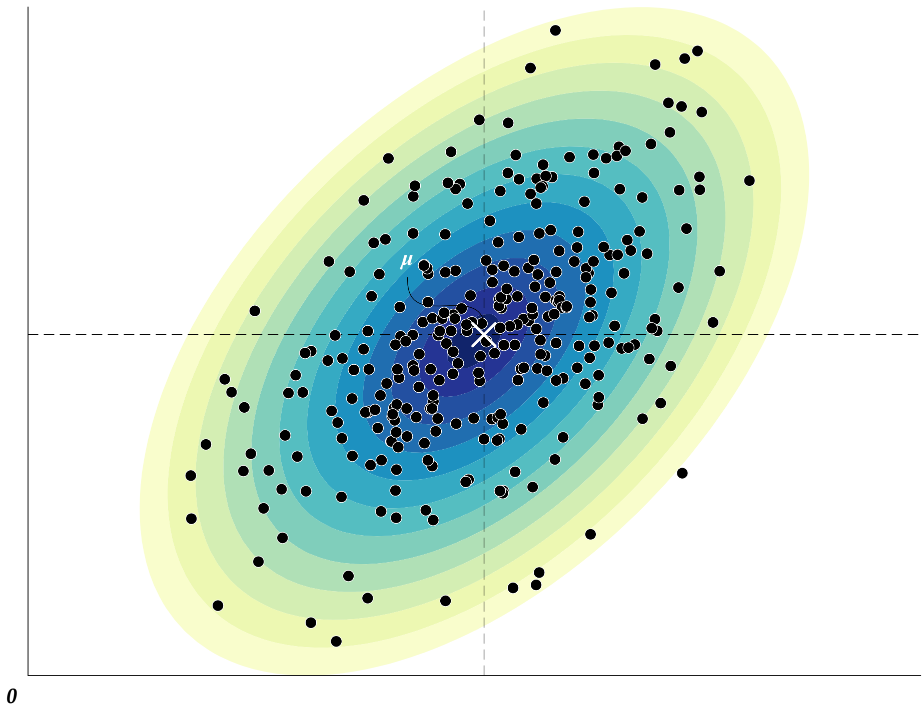


图 2. “数据云”的质心

如图 3 所示，数据矩阵 $\boldsymbol{X}$ 向横轴投影得到 $\boldsymbol{x}_1$，即

$$\boldsymbol{X}@\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}\boldsymbol{x}_1 & \boldsymbol{x}_2\end{bmatrix}@\begin{bmatrix}1\\0\end{bmatrix}=\boldsymbol{x}_1 \tag{3}$$

质心 $\boldsymbol{\mu}$ 向 x_1 轴投影，便得到 μ_1

$$\mu_1=\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{e}_1=\boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{\mu}=\begin{bmatrix}1 & 0\end{bmatrix}@\begin{bmatrix}\mu_1\\\mu_2\end{bmatrix} \tag{4}$$

上式本质上是标量投影。

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。
代码及 PDF 文件下载：https://github.com/Visualize-ML
本书配套微课视频均发布在 B 站——生姜 DrGinger：https://space.bilibili.com/513194466
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

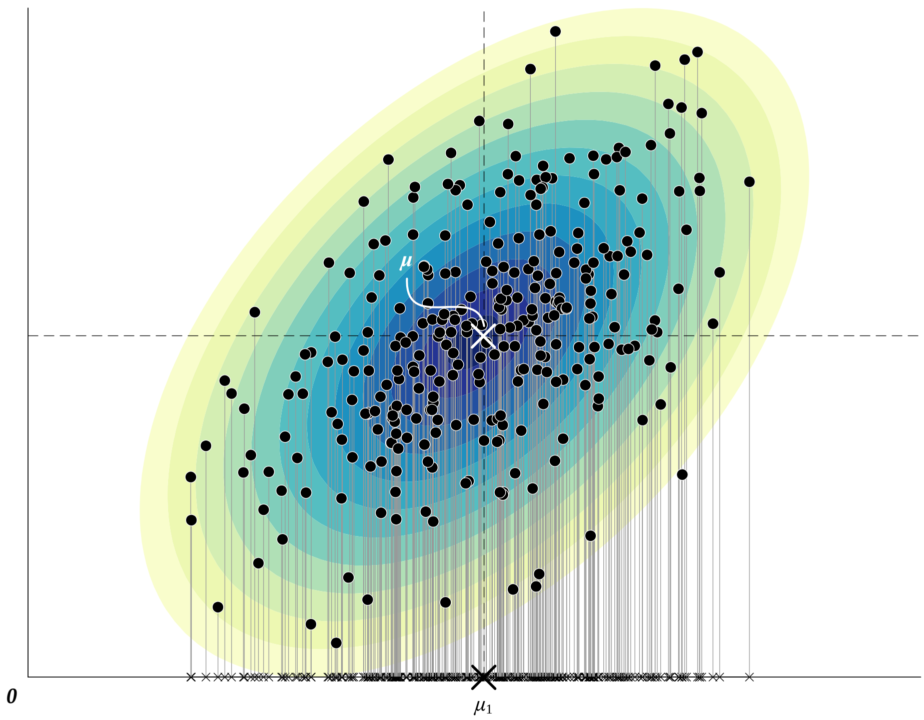


图 3. 数据矩阵向横轴投影

类似地，如图 4 所示，数据矩阵 $\boldsymbol{X}$ 向纵轴投影得到 $\boldsymbol{x}_2$，即

$$\boldsymbol{X} @ \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 \end{bmatrix} @ \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \boldsymbol{x}_2 \tag{5}$$

质心 $\boldsymbol{\mu}$ 向 x_2 轴投影，便得到 μ_2

$$\mu_2 = \boldsymbol{\mu}^{\mathrm{T}} \boldsymbol{e}_2 = \boldsymbol{e}_2^{\mathrm{T}} \boldsymbol{\mu} = \begin{bmatrix} 0 & 1 \end{bmatrix} @ \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \tag{6}$$

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。
代码及 PDF 文件下载：https://github.com/Visualize-ML
本书配套微课视频均发布在 B 站——生姜 DrGinger：https://space.bilibili.com/513194466
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

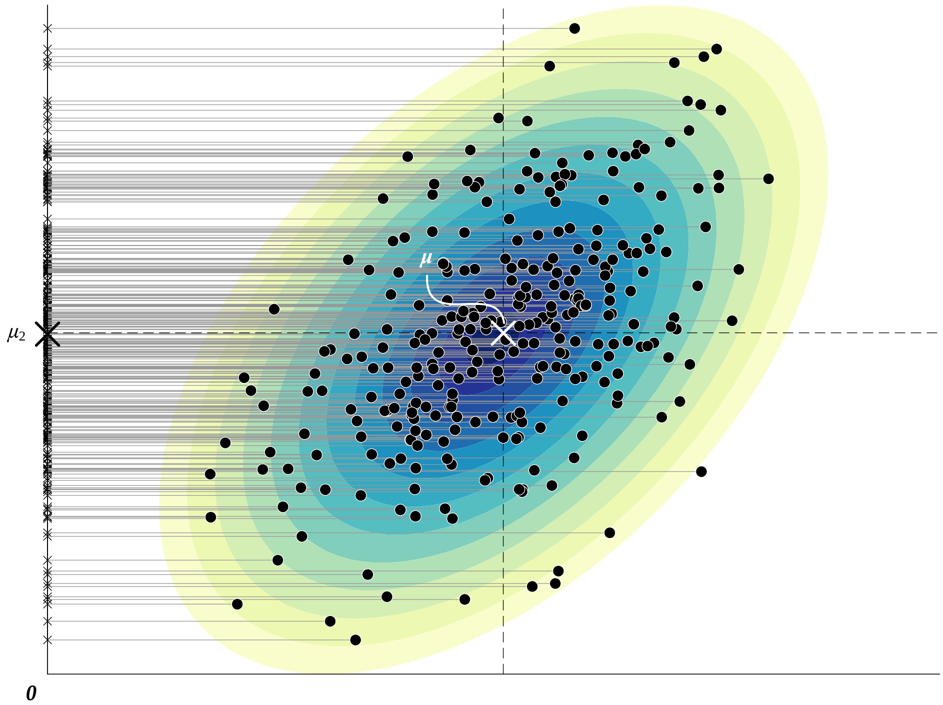


图 4. 数据矩阵向纵轴投影

数据中心化

上一节提过，数据中心化 (去均值) 是指数据矩阵 $\boldsymbol{X}$ 每列减去其所在列的均值得到 $\boldsymbol{X}_{\mathrm{c}}$。用广播原则

$$\boldsymbol{X}_c = \boldsymbol{X} - \boldsymbol{\mu}^{\mathrm{T}} \tag{7}$$

如图 5 所示，数据 (列) 中心化没有改变数据的分布形状，但会将数据中心平移到原点 (零向量 $\boldsymbol{0}$)。

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。
代码及 PDF 文件下载：https://github.com/Visualize-ML
本书配套微课视频均发布在 B 站——生姜 DrGinger：https://space.bilibili.com/513194466
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

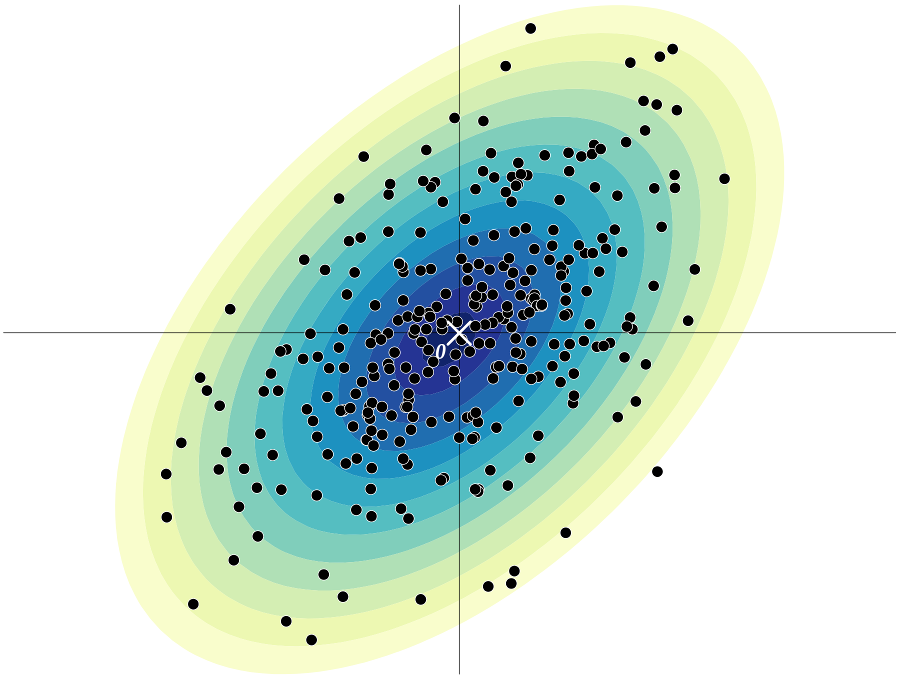


图 5. 数据中心化

协方差矩阵

使用中心化后的数据，计算特征之间的协方差矩阵，表示各变量之间的线性相关性。

上一节提到，协方差矩阵相当于中心化数据矩阵的一种格拉姆矩阵，即

$$\boldsymbol{\Sigma} = \frac{\boldsymbol{X}_c^{\mathrm{T}} \boldsymbol{X}_c}{n-1} \tag{8}$$

带入具体值，我们得到图 1 数据的协方差矩阵

$$\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix} \tag{9}$$

协方差矩阵主对角线为方差，也就是说图 3、图 4 不同轴投影数据的方差都是 1。开平方得到标准差，也都是 1。协方差矩阵非主对角线元素为协方差，这意味着 $\boldsymbol{x}_1$、$\boldsymbol{x}_2$ 的协方差为 0.5。

$\boldsymbol{x}_1$、$\boldsymbol{x}_2$ 的线性相关性系数也是 0.5，即

$$\rho_{1,2} = \frac{\operatorname{cov}\left(\boldsymbol{x}_1, \boldsymbol{x}_2\right)}{\sigma_1 \sigma_2} = \frac{0.5}{1 \times 1} = 0.5 \tag{10}$$

⚠ 注意，由于图 1 数据方差 (标准差) 相差不大，直接对协方差矩阵进行特征值分解完成主成分分析即可；但是，如果方差相差很大，需要先对数据标准化，这是下一节要介绍的话题。

特征值分解

对 (9) 协方差矩阵进行特征值分解

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。
代码及 PDF 文件下载：https://github.com/Visualize-ML
本书配套微课视频均发布在 B 站——生姜 DrGinger：https://space.bilibili.com/513194466
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

$$\boldsymbol{\Sigma}=\begin{bmatrix}1 & 0.5\\ 0.5 & 1\end{bmatrix}=\underbrace{\begin{bmatrix}\sqrt{2}/2 & -\sqrt{2}/2\\ \sqrt{2}/2 & \sqrt{2}/2\end{bmatrix}}_{\boldsymbol{V}}\underbrace{\begin{bmatrix}1.5 & 0\\ 0 & 0.5\end{bmatrix}}_{\boldsymbol{\Lambda}}\underbrace{\begin{bmatrix}\sqrt{2}/2 & \sqrt{2}/2\\ -\sqrt{2}/2 & \sqrt{2}/2\end{bmatrix}}_{\boldsymbol{V}^{\mathrm{T}}}=\boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{\mathrm{T}} \tag{11}$$

上式中 $\boldsymbol{V}$ 和 $\boldsymbol{\Lambda}$ 到底有怎样的含义？

让我们还是用几何视角分析。

如图 6 所示，中心化数据朝 $\boldsymbol{v}$ 方向投影，得到的投影数据 $\boldsymbol{y}$ 为

$$\boldsymbol{y}=\boldsymbol{X}_{\mathrm{c}}\boldsymbol{v} \tag{12}$$

由于 $\boldsymbol{X}_{\mathrm{c}}$ 的质心位于原点，$\boldsymbol{y}$ 的均值为 0。

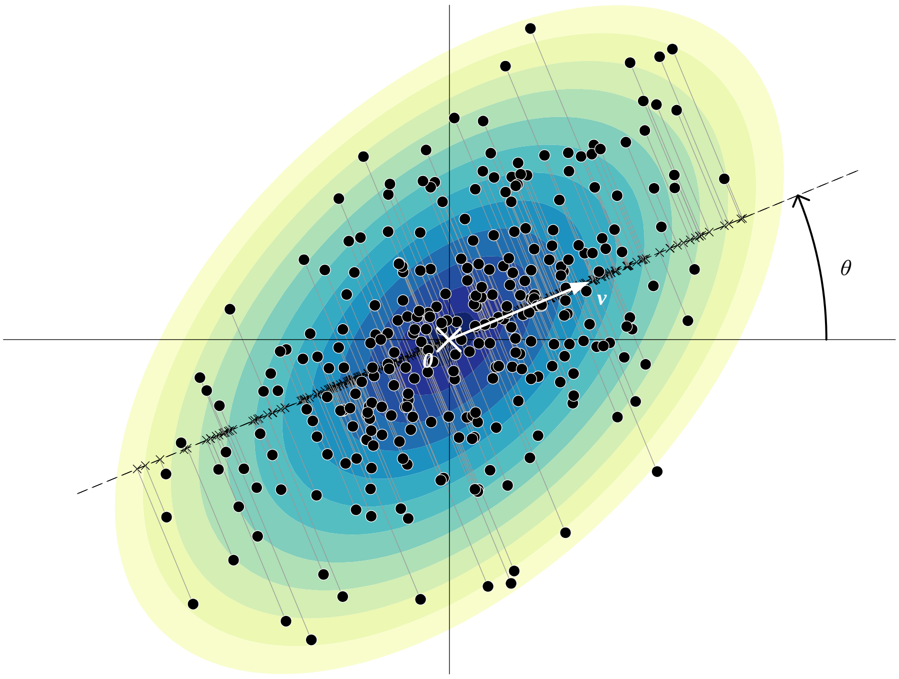


图 6. 中心化数据投影

计算投影数据 $\boldsymbol{y}$ 的方差

$$\operatorname{var}(\boldsymbol{y})=\frac{\boldsymbol{y}^{\mathrm{T}}\boldsymbol{y}}{n-1}=\frac{(\boldsymbol{X}_{\mathrm{c}}\boldsymbol{v})^{\mathrm{T}}\boldsymbol{X}_{\mathrm{c}}\boldsymbol{v}}{n-1}=\boldsymbol{v}^{\mathrm{T}}\frac{\boldsymbol{X}_{\mathrm{c}}^{\mathrm{T}}\boldsymbol{X}_{\mathrm{c}}}{n-1}\boldsymbol{v}=\boldsymbol{v}^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{v} \tag{13}$$

向量 $\boldsymbol{v}$ 和横轴正方向的夹角为 θ。

图 7 所示为方差随 θ 变化，而上式的最大值对应协方差矩阵 $\boldsymbol{\Sigma}$ 的最大特征值 (1.5)，最小值为最小特征值 (0.5)。

(13) 是特殊的瑞利商。此外，大家是否发现图 7 为三角函数？！这些都是本书第 13 章、第 3 节要讨论的话题。

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。
代码及 PDF 文件下载：https://github.com/Visualize-ML
本书配套微课视频均发布在 B 站——生姜 DrGinger：https://space.bilibili.com/513194466
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

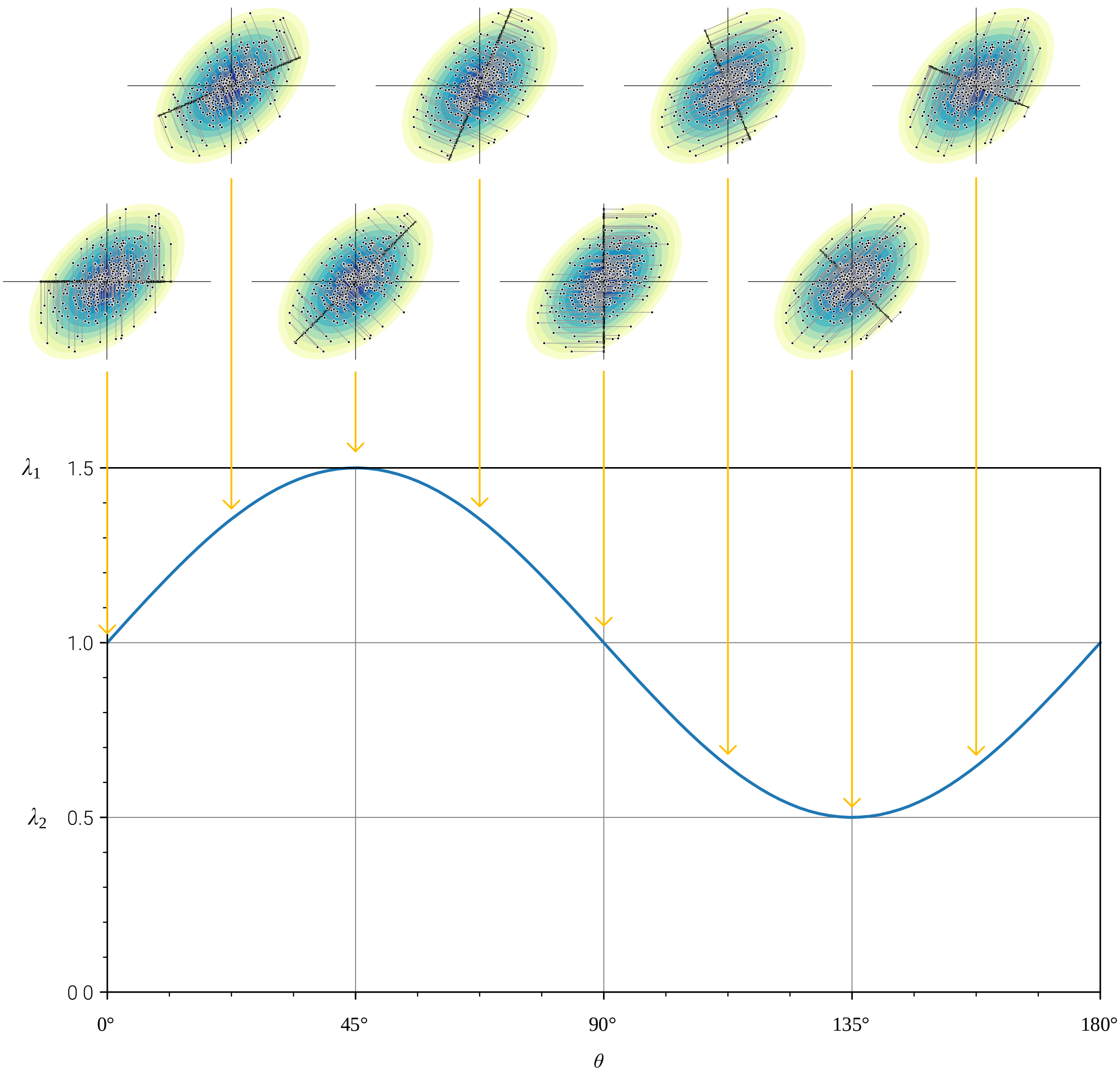


图 7. 方差随 θ 变化

而 $\boldsymbol{V}$ 的两个特征向量就是最大、最小特征值对应的方向，

$$\boldsymbol{V}=\begin{bmatrix}\boldsymbol{v}_1 & \boldsymbol{v}_2\end{bmatrix}=\begin{bmatrix}\begin{bmatrix}\sqrt{2}/2\\ \sqrt{2}/2\end{bmatrix} & \begin{bmatrix}-\sqrt{2}/2\\ \sqrt{2}/2\end{bmatrix}\end{bmatrix} \tag{14}$$

$\boldsymbol{v}_1$、$\boldsymbol{v}_2$ 和横轴正方向夹角分别为 35 度、135 度，对应图 7 最大值、最小值的角度值。

根据 (11)，协方差矩阵可以对角化

$$\boldsymbol{V}^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{V}=\boldsymbol{\Lambda} \tag{15}$$

把 $\boldsymbol{V}$ 写成列向量 $\begin{bmatrix}\boldsymbol{v}_1 & \boldsymbol{v}_2\end{bmatrix}$ 展开上式得到

$$\begin{bmatrix}\boldsymbol{v}_1 & \boldsymbol{v}_2\end{bmatrix}^{\mathrm{T}}\boldsymbol{\Sigma}\begin{bmatrix}\boldsymbol{v}_1 & \boldsymbol{v}_2\end{bmatrix}=\begin{bmatrix}\boldsymbol{v}_1^{\mathrm{T}}\\ \boldsymbol{v}_2^{\mathrm{T}}\end{bmatrix}\boldsymbol{\Sigma}\begin{bmatrix}\boldsymbol{v}_1 & \boldsymbol{v}_2\end{bmatrix}=\begin{bmatrix}\boldsymbol{v}_1^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{v}_1 & \boldsymbol{v}_1^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{v}_2\\ \boldsymbol{v}_2^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{v}_1 & \boldsymbol{v}_2^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{v}_2\end{bmatrix}=\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix} \tag{16}$$

请大家格外注意如下两个等式，我们马上要用到

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。
代码及 PDF 文件下载：https://github.com/Visualize-ML
本书配套微课视频均发布在 B 站——生姜 DrGinger：https://space.bilibili.com/513194466
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

$$
\begin{aligned}
\boldsymbol{v}_1^{\mathrm{T}} \boldsymbol{\Sigma} \boldsymbol{v}_1 &= \lambda_1 \\
\boldsymbol{v}_2^{\mathrm{T}} \boldsymbol{\Sigma} \boldsymbol{v}_2 &= \lambda_2
\end{aligned} \tag{17}
$$

此外，还有一点值得大家注意，协方差的**迹** (trace) 和特征值分解得到的 $\boldsymbol{\Lambda}$ 迹相同，即

$$
\begin{aligned}
\operatorname{trace}(\boldsymbol{\Sigma}) &= \operatorname{trace}\left(\begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}\right) = 1+1=2 \\
\operatorname{trace}(\boldsymbol{\Lambda}) &= \operatorname{trace}\left(\begin{bmatrix} 1.5 & 0 \\ 0 & 0.5 \end{bmatrix}\right) = 1.5+0.5=2
\end{aligned} \tag{18}
$$

简单来说，一个方阵的迹等于它所有特征值 (包括重复) 的总和。

投影

特征值越大，说明该方向上数据的方差越大，信息越丰富。

如图 8 所示，中心化数据矩阵 $\boldsymbol{X}_c$ 向 $\boldsymbol{v}_1$ 投影数据为 $\boldsymbol{y}_1$

$$
\boldsymbol{y}_1 = \boldsymbol{X}_c \boldsymbol{v}_1 \tag{19}
$$

计算 $\boldsymbol{y}_1$ 方差

$$
\operatorname{var}(\boldsymbol{y}_1) = \frac{\boldsymbol{y}_1^{\mathrm{T}} \boldsymbol{y}_1}{n-1} = \frac{(\boldsymbol{X}_c \boldsymbol{v}_1)^{\mathrm{T}} \boldsymbol{X}_c \boldsymbol{v}_1}{n-1} = \boldsymbol{v}_1^{\mathrm{T}} \frac{\boldsymbol{X}_c^{\mathrm{T}} \boldsymbol{X}_c}{n-1} \boldsymbol{v}_1 = \boldsymbol{v}_1^{\mathrm{T}} \boldsymbol{\Sigma} \boldsymbol{v}_1 = \lambda_1 \tag{20}
$$

这意味着 $\boldsymbol{y}_1$ 的方差为 1.5 (λ_1)。

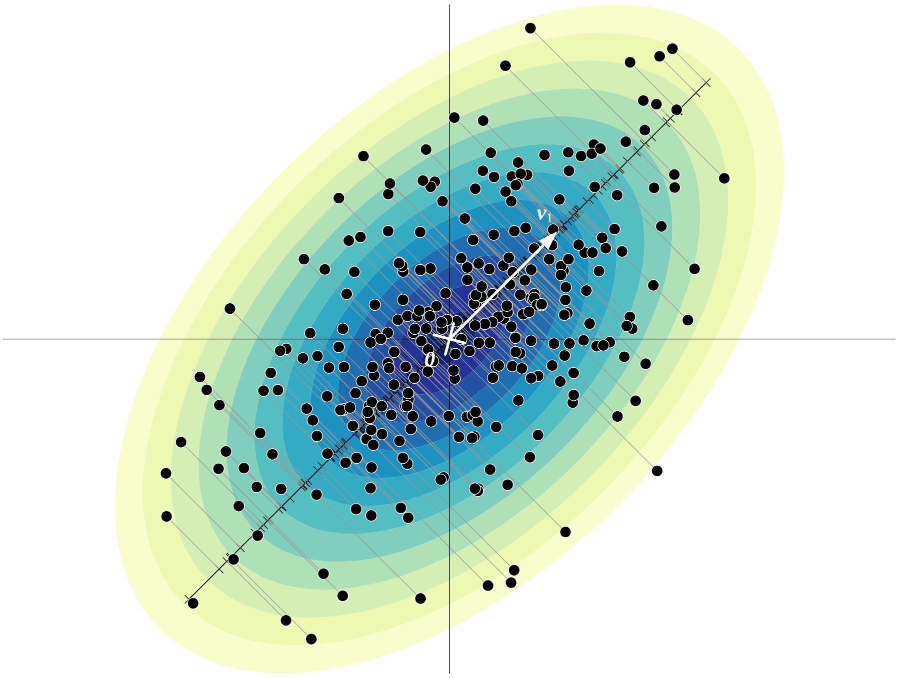


图 8. 数据朝 $\boldsymbol{v}_1$ 投影

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。
代码及 PDF 文件下载：https://github.com/Visualize-ML
本书配套微课视频均发布在 B 站——生姜 DrGinger：https://space.bilibili.com/513194466
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

类似地，如图 9 所示，中心化数据矩阵 $\boldsymbol{X}_c$ 向 $\boldsymbol{v}_2$ 投影数据为 $\boldsymbol{y}_2$

$$\boldsymbol{y}_2 = \boldsymbol{X}_c \boldsymbol{v}_2 \tag{21}$$

计算 $\boldsymbol{y}_2$ 方差

$$\operatorname{var}\left(\boldsymbol{y}_2\right) = \frac{\boldsymbol{y}_2^{\mathrm{T}} \boldsymbol{y}_2}{n-1} = \frac{\left(\boldsymbol{X}_c \boldsymbol{v}_2\right)^{\mathrm{T}} \boldsymbol{X}_c \boldsymbol{v}_2}{n-1} = \boldsymbol{v}_2^{\mathrm{T}} \frac{\boldsymbol{X}_c^{\mathrm{T}} \boldsymbol{X}_c}{n-1} \boldsymbol{v}_2 = \boldsymbol{v}_2^{\mathrm{T}} \boldsymbol{\Sigma} \boldsymbol{v}_2 = \lambda_2 \tag{22}$$

这意味着 $\boldsymbol{y}_2$ 的方差为 0.5 (λ_2)。

观察图 8、图 9，我们发现 $\boldsymbol{v}_1$ 对应椭圆的长轴方向，也是第一主元的方向；$\boldsymbol{v}_2$ 对应椭圆的短轴方向，也是第二主元的方向。

数据朝不同方向投影会得到不同的投影结果，对应不同的分布；朝椭圆长轴方向投影，得到的数据标准差最大；朝椭圆短轴方向投影得到的数据标准差最小。

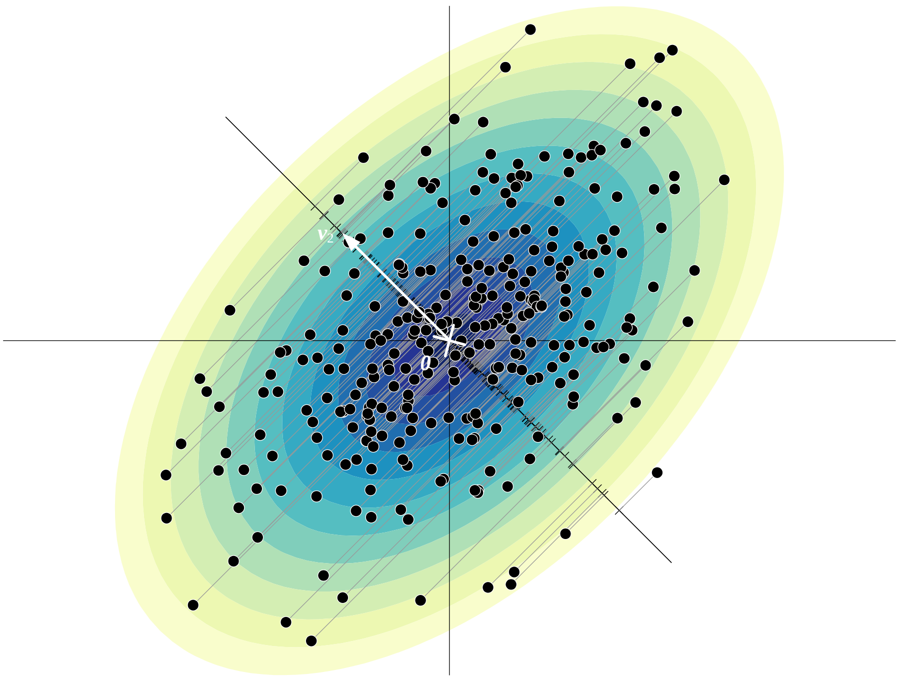


图 9. 数据朝 $\boldsymbol{v}_2$ 投影

在规范正交基 V 中看数据

由于协方差矩阵为对阵矩阵，因此特征值分解 (谱分解) 得到的 $\boldsymbol{V}$ 为正交矩阵。$\boldsymbol{V}$ 的列向量构成规范正交基。

换个视角来看，如图 10 所示，主成分分析无非就是 $\boldsymbol{V}$ 中看同一组数据。

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。
代码及 PDF 文件下载：https://github.com/Visualize-ML
本书配套微课视频均发布在 B 站——生姜 DrGinger：https://space.bilibili.com/513194466
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

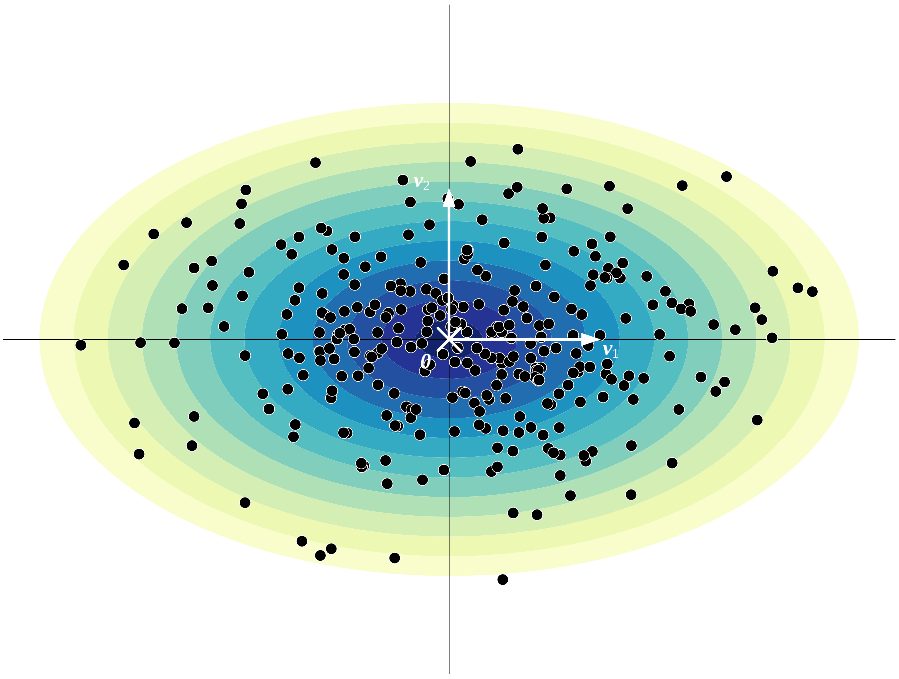


图 10. 在 $[v_1, v_2]$ 中看中心化数据

图 11 所示为在 $[v_1, v_2]$ 中朝 v_1 投影结果。图 12 所示为在 $[v_1, v_2]$ 中朝 v_2 投影结果。

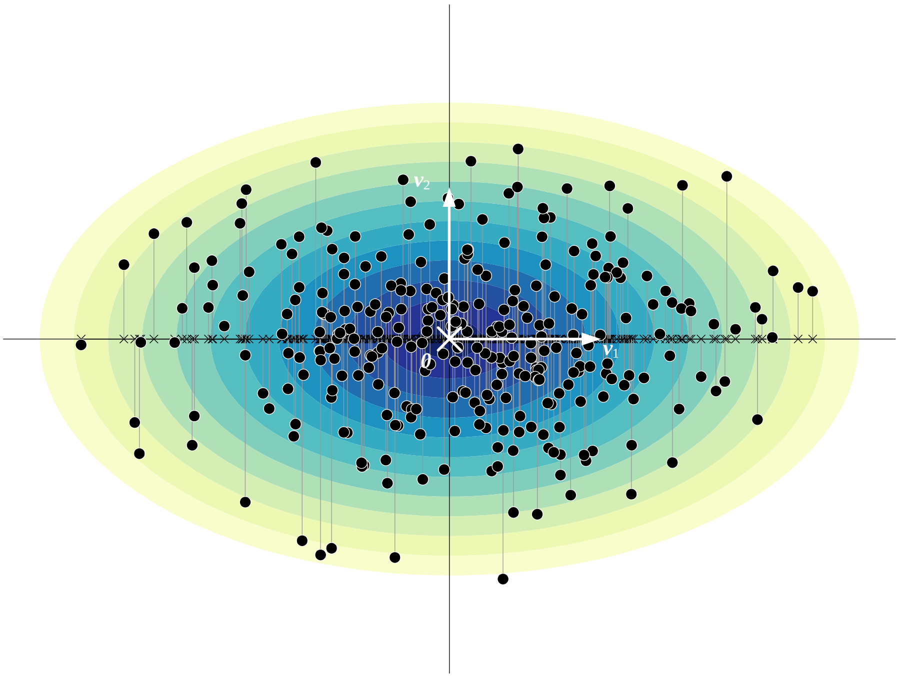


图 11. 在 $[v_1, v_2]$ 中看中心化数据，朝 v_1 投影

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。
代码及 PDF 文件下载：https://github.com/Visualize-ML
本书配套微课视频均发布在 B 站——生姜 DrGinger：https://space.bilibili.com/513194466
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

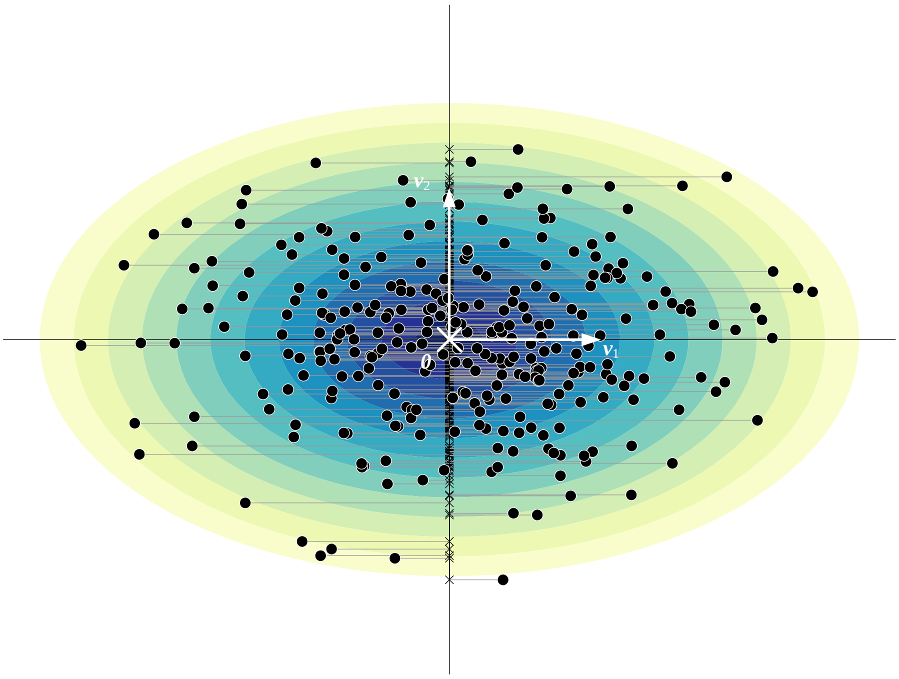


图 12. 在 [v_1, v_2] 中看中心化数据，朝 v_2 投影

LA_12_02_01.ipynb 完成本节主要运算以及可视化，请大家自学。

请大家用 DeepSeek/ChatGPT 等工具完成本节如下习题。

Q1. 请大家了解马氏距离。

Q2. 请大家自学 LA_12_02_01.ipynb 中的 mahalanobis_distance() 这个计算马氏距离的函数。

Q3. 请大家将 LA_12_02_01.ipynb 可视化函数写成一个自定义函数，方便反复调用。

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。
版权归清华大学出版社所有，请勿商用，引用请注明出处。
代码及 PDF 文件下载：https://github.com/Visualize-ML
本书配套微课视频均发布在 B 站——生姜 DrGinger：https://space.bilibili.com/513194466
欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com