20 Revisit Conic Sections 再谈圆锥曲线

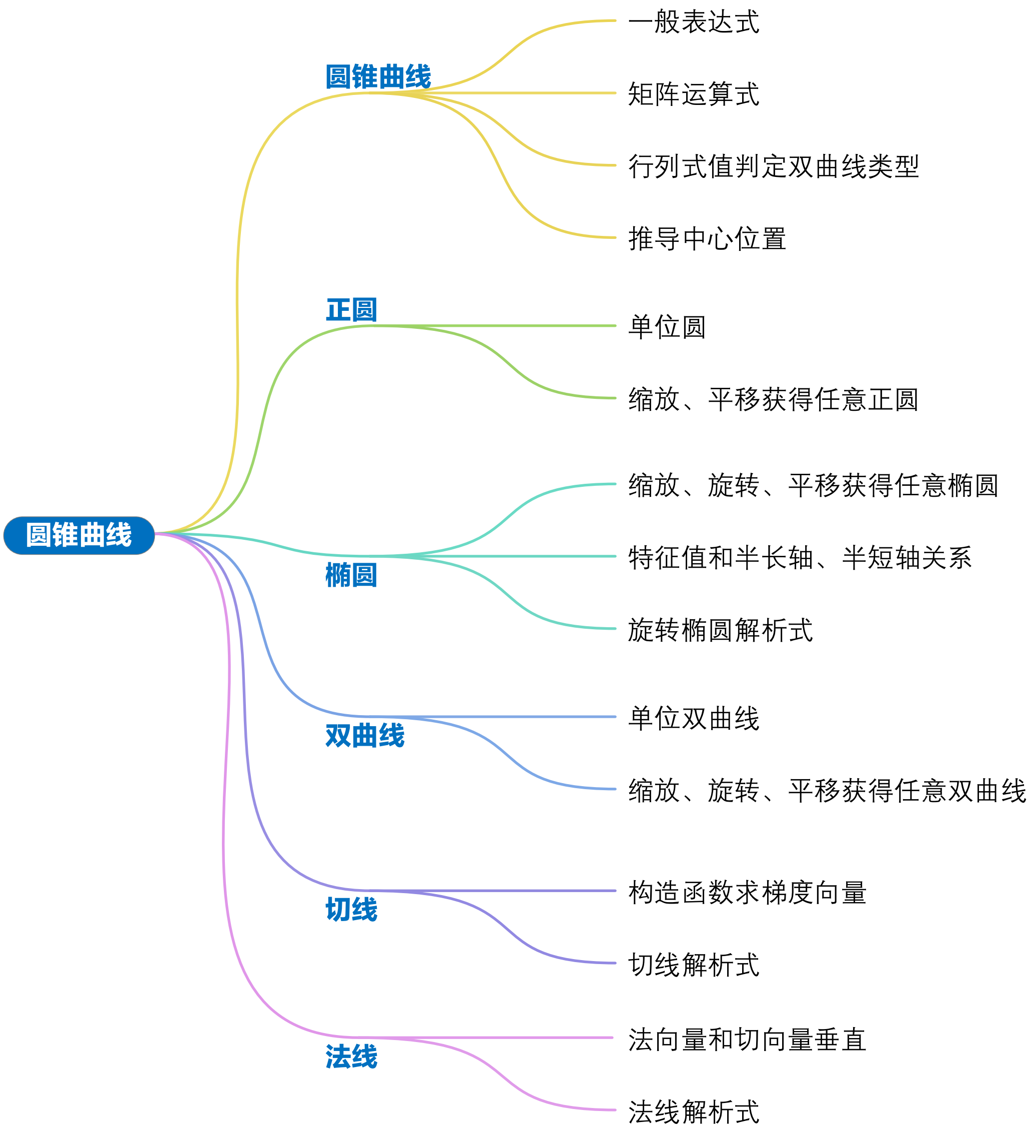
从矩阵运算和几何变换视角

滴水穿石，靠的不是力量，而是坚持。

Dripping water hollows out stone, not through force but through persistence.

—— 奥维德 (Ovid) | 古罗马诗人 | 43 BC ~ 17/18 AD

- ◄ `matplotlib.pyplot.contour()` 绘制等高线图
- ◄ `numpy.cos()` 计算余弦值
- ◄ `numpy.linalg.eig()` 特征值分解
- ◄ `numpy.linalg.inv()` 矩阵求逆
- ◄ `numpy.sin()` 计算正弦值
- ◄ `numpy.tan()` 计算正切


圆锥曲线
圆锥曲线
一般表达式
矩阵运算式
行列式值判定双曲线类型
推导中心位置
正圆
单位圆
缩放、平移获得任意正圆
椭圆
缩放、旋转、平移获得任意椭圆
特征值和半长轴、半短轴关系
旋转椭圆解析式
双曲线
单位双曲线
缩放、旋转、平移获得任意双曲线
切线
构造函数求梯度向量
切线解析式
法线
法向量和切向量垂直
法线解析式

20.1 无处不在的圆锥曲线

本套丛书每一册几乎都离不开圆锥曲线这个话题。

《数学要素》一本详细介绍过圆锥曲线相关性质，《概率统计》一册讨论圆锥曲线和高斯分布千丝万缕的联系，同时我们也看到条件概率、回归分析和主成分分析中圆锥曲线扮演重要角色。《机器学习》一册介绍的很多算法中，决策边界就是圆锥曲线。

利用本书前文讲解的线性代数工具，本章将从矩阵运算和几何变换角度探讨圆锥曲线。这个视角将帮助大家更加深刻理解圆锥曲线在概率统计、数据学科和机器学习的重要作用。

一般表达式

圆锥曲线一般表达式，如下：

$$Ax_1^2 + Bx_1x_2 + Cx_2^2 + Dx_1 + Ex_2 + F = 0 \tag{1}$$

把 (1) 写成矩阵运算式：

$$\frac{1}{2}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} 2A & B \\ B & 2C \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} D \\ E \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + F = 0 \tag{2}$$

(2) 进一步写成：

$$\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x} + \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} + F = 0 \tag{3}$$

其中，

$$\boldsymbol{Q} = \begin{bmatrix} 2A & B \\ B & 2C \end{bmatrix}, \quad \boldsymbol{w} = \begin{bmatrix} D \\ E \end{bmatrix} \tag{4}$$

矩阵 $\boldsymbol{Q}$ 的行列式值为：

$$\det \boldsymbol{Q} = \det \begin{bmatrix} 2A & B \\ B & 2C \end{bmatrix} = 4AC - B^2 \tag{5}$$

矩阵 $\boldsymbol{Q}$ 的行列式值决定了圆锥曲线的形状：

- 当 $4AC - B^2 > 0$，上式为椭圆；特别地，当 $A = C$ 且 $B = 0$，解析式为正圆；
- $4AC - B^2 = 0$ 时，解析式为抛物线；
- $4AC - B^2 < 0$ 时，解析式为双曲线。

这实际上回答了本系列丛书《数学要素》中提出的一个问题。

中心

当 $4AC - B^2$ 不等于 0，圆锥曲线 (椭圆、正圆和双曲线) 存在其中心。依照 (2) 形式构造如下二元函数 $f(x_1, x_2)$：

$$f(x_1,x_2)=\frac{1}{2}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^{\mathrm{T}}\begin{bmatrix} 2A & B \\ B & 2C \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}+\begin{bmatrix} D \\ E \end{bmatrix}^{\mathrm{T}}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}+F \tag{6}$$

下面介绍如何求解圆锥曲线中心。

$f(x_1, x_2)$ 对 $[x_1, x_2]^{\mathrm{T}}$ 一阶导数为 $[0, 0]^{\mathrm{T}}$ 时，(x_1, x_2) 为 $f(x_1, x_2)$ 驻点，即圆锥曲线中心：

$$\frac{\partial f}{\partial \boldsymbol{x}}=\begin{bmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{7}$$

推导圆锥曲线中心位置：

$$\begin{bmatrix} 2A & B \\ B & 2C \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}+\begin{bmatrix} D \\ E \end{bmatrix}=\boldsymbol{0} \quad \Rightarrow \quad \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=-\begin{bmatrix} 2A & B \\ B & 2C \end{bmatrix}^{-1}\begin{bmatrix} D \\ E \end{bmatrix} \tag{8}$$

得到圆锥曲线中心坐标：

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}=\frac{1}{B^2-4AC}\begin{bmatrix} 2CD-BE \\ 2AE-DB \end{bmatrix} \tag{9}$$

20.2 正圆：从单位圆到任意正圆

单位圆

平面上，圆心位于原点的单位圆可以写成如下形式：

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}-1=0 \tag{10}$$

其中 $\boldsymbol{x}$ 为，

$$\boldsymbol{x}=\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{11}$$

展开 (10) 得到：

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}-1=x_1^2+x_2^2-1=0 \tag{12}$$

当然，(10) 可以用 L^2 范数、向量内积等方式表达，比如：

$$\begin{gathered}\|\boldsymbol{x}\|_2^2-1=0\\ \boldsymbol{x}\cdot\boldsymbol{x}-1=0\\ \langle\boldsymbol{x},\boldsymbol{x}\rangle-1=0\end{gathered}\tag{13}$$

缩放

圆心位于原点半径为 r 的正圆解析式为：

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}-r^2=0\tag{14}$$

上式相当于，

$$\boldsymbol{x}^{\mathrm{T}}\begin{bmatrix}1/r^2 & 0\\ 0 & 1/r^2\end{bmatrix}\boldsymbol{x}-1=0\tag{15}$$

从几何变换视角来观察，相信大家已经在 (15) 中看到起到缩放作用的对角方阵。

缩放 + 平移

圆心位于 $\boldsymbol{c}=[c_1, c_2]^{\mathrm{T}}$ 半径为 r 的正圆解析式为：

$$(\boldsymbol{x}-\boldsymbol{c})^{\mathrm{T}}\begin{bmatrix}1/r^2 & 0\\ 0 & 1/r^2\end{bmatrix}(\boldsymbol{x}-\boldsymbol{c})-1=0\tag{16}$$

(16) 也可以写成：

$$\begin{gathered}(\boldsymbol{x}-\boldsymbol{c})^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{c})-r^2=0\\ \|\boldsymbol{x}-\boldsymbol{c}\|_2^2-r^2=0\\ (\boldsymbol{x}-\boldsymbol{c})\cdot(\boldsymbol{x}-\boldsymbol{c})-r^2=0\\ \langle(\boldsymbol{x}-\boldsymbol{c}),(\boldsymbol{x}-\boldsymbol{c})\rangle-r^2=0\end{gathered}\tag{17}$$

不同参考资料中圆锥曲线的表达各有不同。本节不厌其烦地罗列圆锥曲线的各种形式解析式目的只有一个，让大家对它们感觉不陌生。

为了让大家看到我们熟悉的正圆解析式，进一步展开整理 (16) 得到：

$$\begin{aligned}(\boldsymbol{x}-\boldsymbol{c})^{\mathrm{T}}\begin{bmatrix}1/r^2 & 0\\ 0 & 1/r^2\end{bmatrix}(\boldsymbol{x}-\boldsymbol{c})&=\left(\begin{bmatrix}x_1\\ x_2\end{bmatrix}-\begin{bmatrix}c_1\\ c_2\end{bmatrix}\right)^{\mathrm{T}}\begin{bmatrix}1/r^2 & 0\\ 0 & 1/r^2\end{bmatrix}\left(\begin{bmatrix}x_1\\ x_2\end{bmatrix}-\begin{bmatrix}c_1\\ c_2\end{bmatrix}\right)\\ &=\begin{bmatrix}x_1-c_1 & x_2-c_2\end{bmatrix}\begin{bmatrix}1/r^2 & 0\\ 0 & 1/r^2\end{bmatrix}\begin{bmatrix}x_1-c_1\\ x_2-c_2\end{bmatrix}\\ &=\frac{(x_1-c_2)^2}{r^2}+\frac{(x_1-c_2)^2}{r^2}=1\end{aligned}\tag{18}$$

从单位圆到一般正圆

在 (16) 中，大家应该看到了平移。

下面探讨圆心位于原点的单位圆如何一步步经过几何变换得到 (18) 中对应的圆心位于 $\boldsymbol{c} = [c_1, c_2]^{\mathrm{T}}$ 半径为 r 的正圆。

平面内，单位圆解析式写成：

$$\boldsymbol{z}^{\mathrm{T}}\boldsymbol{z}-1=0 \tag{19}$$

$\boldsymbol{z}$ 通过先等比例缩放，再平移得到 $\boldsymbol{x}$：

$$\boldsymbol{x}=\underbrace{\begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}}_{\text{Scale}}\boldsymbol{z}+\underbrace{\boldsymbol{c}}_{\text{Translate}} \tag{20}$$

$\boldsymbol{z}$ 可以写作：

$$\boldsymbol{z}=\begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}^{-1}(\boldsymbol{x}-\boldsymbol{c}) \tag{21}$$

将 (21) 代入 (19)，得到：

$$\left(\begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}^{-1}(\boldsymbol{x}-\boldsymbol{c})\right)^{\mathrm{T}}\left(\begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}^{-1}(\boldsymbol{x}-\boldsymbol{c})\right)-1=0 \tag{22}$$

整理上式，

$$(\boldsymbol{x}-\boldsymbol{c})^{\mathrm{T}}\begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}^{-1}\begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}^{-1}(\boldsymbol{x}-\boldsymbol{c})-1=0 \tag{23}$$

即，

$$(\boldsymbol{x}-\boldsymbol{c})^{\mathrm{T}}\begin{bmatrix} 1/r^2 & 0 \\ 0 & 1/r^2 \end{bmatrix}(\boldsymbol{x}-\boldsymbol{c})-1=0 \tag{24}$$

可以发现 (24) 和 (16) 完全一致。也就是说，如图 1 所示，单位圆可以通过先缩放再平移，得到圆心位于 $\boldsymbol{c}$ 半径为 r 的圆。

沿着这一思路，下一节我们讨论如何通过一步步几何变换操作将单位圆转换成任意旋转椭圆。

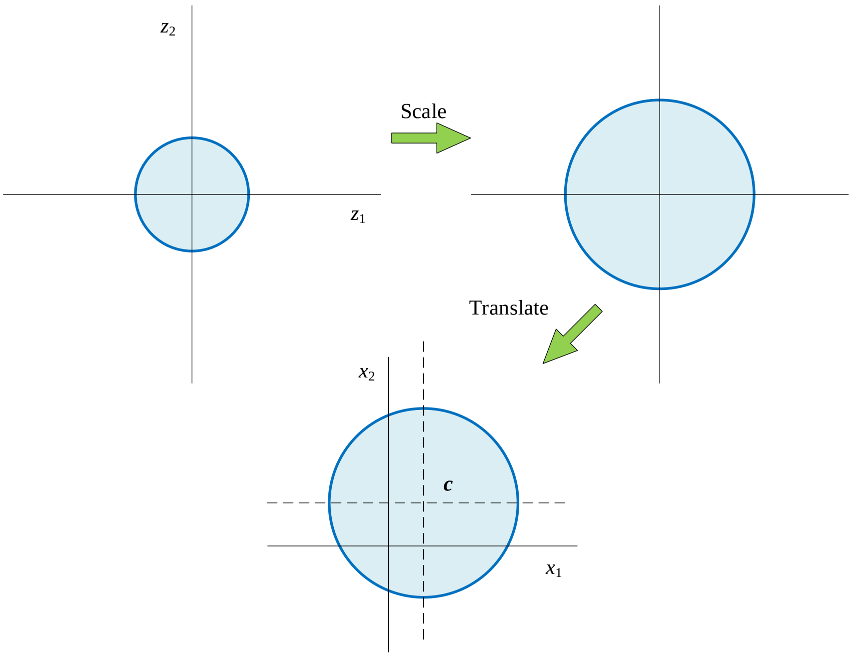


图 1. 单位圆变换得到圆心位于 $\boldsymbol{c}$ 半径为 r 的正圆

20.3 单位圆到旋转椭圆：缩放 → 旋转 → 平移

这一节介绍如何利用"缩放 → 旋转 → 平移" 几何变换，将单位圆变成中心位于任何位置的旋转椭圆。

利用上一节 (19) 给出的单位圆解析式中 $\boldsymbol{z}$，对 $\boldsymbol{z}$ 先用 $\boldsymbol{S}$ 缩放，再通过 $\boldsymbol{R}$ 逆时针旋转 θ，最后平移 $\boldsymbol{c}$，得到 $\boldsymbol{x}$：

$$\underbrace{\boldsymbol{R}}_{\text{Rotate}}\underbrace{\boldsymbol{S}}_{\text{Scale}}\boldsymbol{z}+\underbrace{\boldsymbol{c}}_{\text{Translate}}=\boldsymbol{x} \tag{25}$$

其中，

$$\boldsymbol{R}=\underbrace{\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}}_{\text{Rotate}},\quad \boldsymbol{S}=\underbrace{\begin{bmatrix}a & 0\\ 0 & b\end{bmatrix}}_{\text{Scale}},\quad \boldsymbol{c}=\underbrace{\begin{bmatrix}c_1\\ c_2\end{bmatrix}}_{\text{Translate}} \tag{26}$$

将 (26) 代入 (25)，得到：

$$\underbrace{\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}}_{\text{Rotate}}\underbrace{\begin{bmatrix}a & 0\\ 0 & b\end{bmatrix}}_{\text{Scale}}\begin{bmatrix}z_1\\ z_2\end{bmatrix}+\underbrace{\begin{bmatrix}c_1\\ c_2\end{bmatrix}}_{\text{Translate}}=\begin{bmatrix}x_1\\ x_2\end{bmatrix} \tag{27}$$

图 2 所示为从单位圆经过"缩放 → 旋转 → 平移"几何变换得到中心位于 $\boldsymbol{c}$ 的旋转椭圆过程。

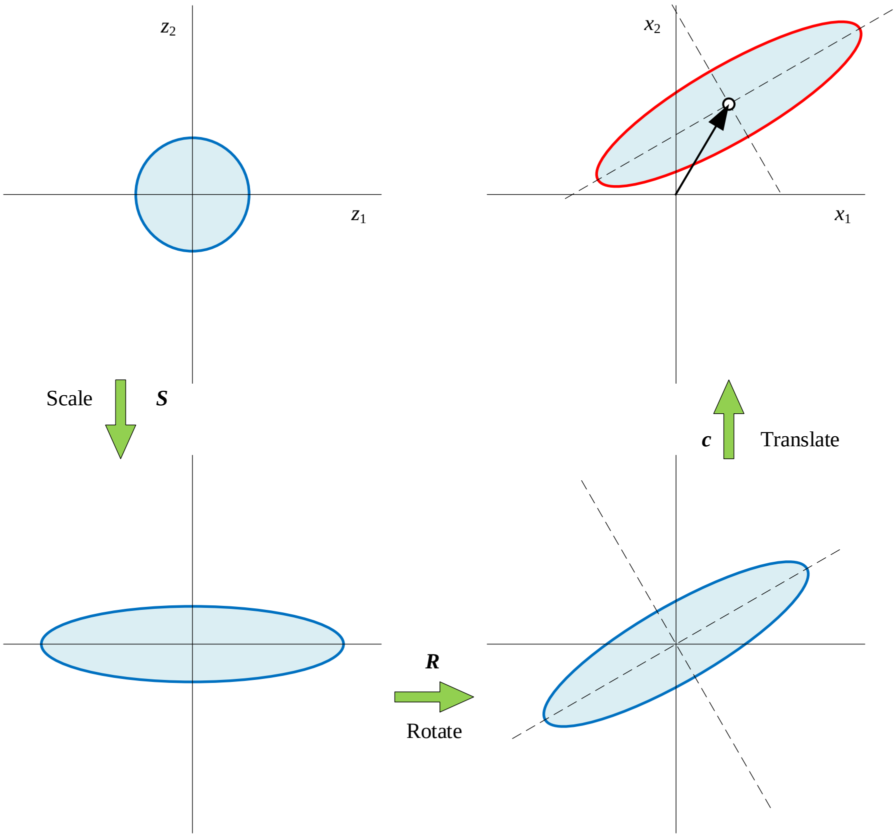


图 2. 从单位圆得到旋转椭圆

整理 (25) 得到 z 的解析式为:

$$z = S^{-1}R^{-1}(x - c) \tag{28}$$

R 为正交矩阵, 所以:

$$R^{-1} = R^{\mathrm{T}} \tag{29}$$

(28) 写成:

$$z = S^{-1}R^{\mathrm{T}}(x - c) \tag{30}$$

将 (30) 代入 (19) 正圆解析式, 得到:

$$\left(S^{-1}R^{\mathrm{T}}(x - c)\right)^{\mathrm{T}}\left(S^{-1}R^{\mathrm{T}}(x - c)\right) - 1 = 0 \tag{31}$$

进一步整理得到:

$$(x - c)^{\mathrm{T}} RS^{-2}R^{\mathrm{T}}(x - c) - 1 = 0 \tag{32}$$

令

$$Q = RS^{-2}R^{\mathrm{T}} = R\begin{bmatrix} a^{-2} & \\ & b^{-2} \end{bmatrix}R^{\mathrm{T}} \tag{33}$$

对 $\boldsymbol{Q}$ 特征值分解，得到：

$$Q = R\begin{bmatrix} \lambda_1 & \\ & \lambda_2 \end{bmatrix}R^{\mathrm{T}} \tag{34}$$

比较 (33) 和 (34)，得出特征值矩阵和缩放矩阵关系：

$$\begin{bmatrix} \lambda_1 & \\ & \lambda_2 \end{bmatrix} = \begin{bmatrix} a^{-2} & \\ & b^{-2} \end{bmatrix} \tag{35}$$

即，

$$a = \frac{1}{\sqrt{\lambda_1}}, \quad b = \frac{1}{\sqrt{\lambda_2}} \tag{36}$$

这样，我们便得到椭圆半长轴和半短轴长度和矩阵 $\boldsymbol{Q}$ 特征值之间的关系。

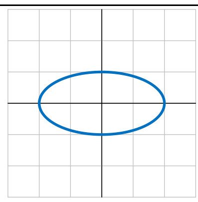
此外，反方向来看图 2，也就是说中心位于任何位置的旋转椭圆可以通过“平移 → 旋转 → 缩放”变成单位圆。注意单位圆默认圆心位于原点。

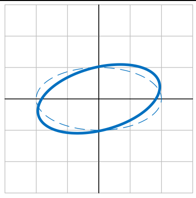
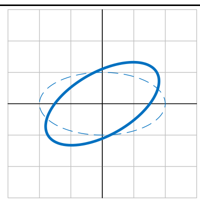
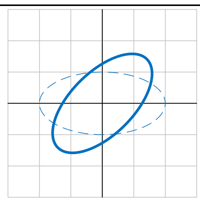
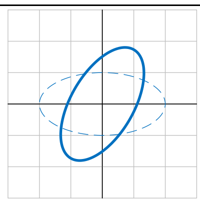
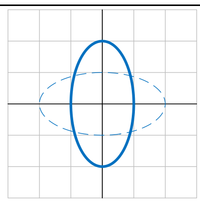
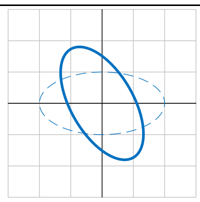
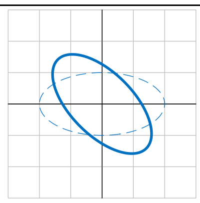
只考虑旋转

表 1 中给出一系列不同旋转角度椭圆解析式和对应 $\boldsymbol{Q}$ 的特征值分解。表 1 中椭圆半长轴和半短轴长度保持一致，唯一变化的就是旋转角度。对几组 $\boldsymbol{Q}$ 特征值分解发现特征值完全相同。

表中特征值在对角线方向从小到大排列，它们的平方根倒数分别对应椭圆的半长轴和半短轴长度。

表 1. 旋转椭圆解析式、$\boldsymbol{Q}$ 的特征值分解

旋转角度	椭圆解析式 (最多保留小数点后 4 位)	对 $\boldsymbol{Q}$ 特征值分解 (最多保留小数点后 4 位)	图像
$\theta = 0°$ (0)	$x^{\mathrm{T}}\begin{bmatrix} 0.25 & 0 \\ 0 & 1 \end{bmatrix}x - 1 = 0$	$\begin{bmatrix} 0.25 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0.25 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	

$\theta=15°$ ($\pi/12$)	$\boldsymbol{x}^{\mathrm{T}}\begin{bmatrix}0.3002 & -0.1875\\ -0.1875 & 0.9498\end{bmatrix}\boldsymbol{x}-1=0$	$\begin{bmatrix}0.3002 & -0.1875\\ -0.1875 & 0.9498\end{bmatrix}=\begin{bmatrix}0.9659 & -0.2588\\ 0.2588 & 0.9659\end{bmatrix}\begin{bmatrix}0.25 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}0.9659 & 0.2588\\ -0.2588 & 0.9659\end{bmatrix}$	
$\theta=30°$ ($\pi/6$)	$\boldsymbol{x}^{\mathrm{T}}\begin{bmatrix}0.4375 & -0.3248\\ -0.3248 & 0.8125\end{bmatrix}\boldsymbol{x}-1=0$	$\begin{bmatrix}0.4375 & -0.3248\\ -0.3248 & 0.8125\end{bmatrix}=\begin{bmatrix}0.8660 & -0.5000\\ 0.5000 & 0.8660\end{bmatrix}\begin{bmatrix}0.25 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}0.8660 & 0.5000\\ -0.5000 & 0.8660\end{bmatrix}$	
$\theta=45°$ ($\pi/4$)	$\boldsymbol{x}^{\mathrm{T}}\begin{bmatrix}0.6250 & -0.3750\\ -0.3750 & 0.6250\end{bmatrix}\boldsymbol{x}-1=0$	$\begin{bmatrix}0.6250 & -0.3750\\ -0.3750 & 0.6250\end{bmatrix}=\begin{bmatrix}0.7071 & -0.7071\\ 0.7071 & 0.7071\end{bmatrix}\begin{bmatrix}0.25 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}0.7071 & 0.7071\\ -0.7071 & 0.7071\end{bmatrix}$	
$\theta=60°$ ($\pi/3$)	$\boldsymbol{x}^{\mathrm{T}}\begin{bmatrix}0.8125 & -0.3248\\ -0.3248 & 0.4375\end{bmatrix}\boldsymbol{x}-1=0$	$\begin{bmatrix}0.8125 & -0.3248\\ -0.3248 & 0.4375\end{bmatrix}=\begin{bmatrix}0.5000 & -0.8660\\ 0.8660 & 0.5000\end{bmatrix}\begin{bmatrix}0.25 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}0.5000 & 0.8660\\ -0.8660 & 0.5000\end{bmatrix}$	
$\theta=90°$ ($\pi/2$)	$\boldsymbol{x}^{\mathrm{T}}\begin{bmatrix}1 & 0\\ 0 & 0.25\end{bmatrix}\boldsymbol{x}-1=0$	$\begin{bmatrix}1 & 0\\ 0 & 0.25\end{bmatrix}=\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\begin{bmatrix}0.25 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}$	
$\theta=120°$ ($2\pi/3$)	$\boldsymbol{x}^{\mathrm{T}}\begin{bmatrix}0.8125 & 0.3248\\ 0.3248 & 0.4375\end{bmatrix}\boldsymbol{x}-1=0$	$\begin{bmatrix}0.8125 & 0.3248\\ 0.3248 & 0.4375\end{bmatrix}=\begin{bmatrix}-0.5000 & -0.8660\\ 0.8660 & -0.5000\end{bmatrix}\begin{bmatrix}0.25 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}-0.5000 & 0.8660\\ -0.8660 & -0.5000\end{bmatrix}$	
$\theta=145°$ ($3\pi/4$)	$\boldsymbol{x}^{\mathrm{T}}\begin{bmatrix}0.4967 & 0.3524\\ 0.3524 & 0.7533\end{bmatrix}\boldsymbol{x}-1=0$	$\begin{bmatrix}0.4967 & 0.3524\\ 0.3524 & 0.7533\end{bmatrix}=\begin{bmatrix}-0.8192 & -0.5736\\ 0.5736 & -0.8192\end{bmatrix}\begin{bmatrix}0.25 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}-0.8192 & 0.5736\\ -0.5736 & -0.8192\end{bmatrix}$	

一般解析式

为了方便整理旋转椭圆解析式，省略 (30) 平移项 $\boldsymbol{c}$，将其展开，得到：

$$
\begin{aligned}
\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \boldsymbol{S}^{-1}\boldsymbol{R}^{\mathrm{T}} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}^{-1} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\
&= \begin{bmatrix} 1/a & 0 \\ 0 & 1/b \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\
&= \begin{bmatrix} \dfrac{\cos\theta}{a} x_1 + \dfrac{\sin\theta}{a} x_2 \\ -\dfrac{\sin\theta}{b} x_1 + \dfrac{\cos\theta}{b} x_2 \end{bmatrix}
\end{aligned} \tag{37}
$$

将 (37) 代入 (19)，整理得到旋转椭圆解析式：

$$
\frac{\left[x_1 \cos(\theta) + x_2 \sin(\theta)\right]^2}{a^2} + \frac{\left[x_1 \sin(\theta) - x_2 \cos(\theta)\right]^2}{b^2} = 1 \tag{38}
$$

这和《数学要素》给出的旋转椭圆解析式完全一致。图 3 所示为单位圆经过几何变换获得中心位于 (2, 1) 的几个不同旋转角度椭圆的示例。

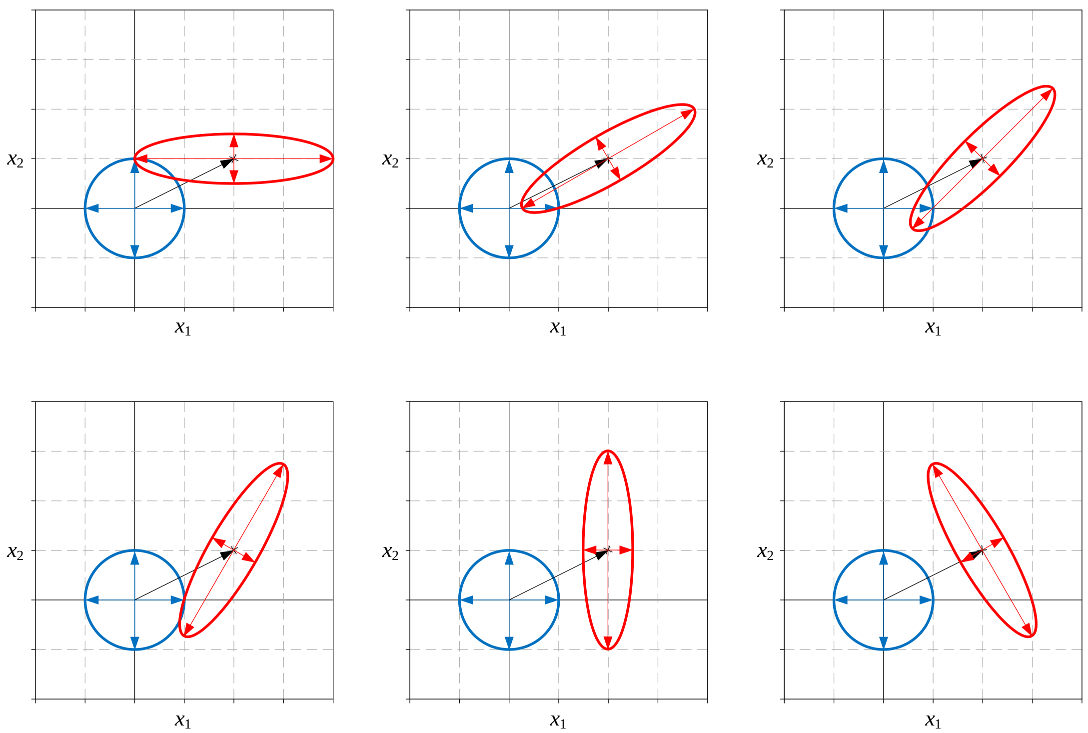


图 3. 通过单位圆获得几个不同的旋转椭圆

Bk4_Ch20_01.py 绘制图 3。

```python
# Bk4_Ch20_01.py

import numpy as np
import matplotlib.pyplot as plt

alphas = np.linspace(0, 2*np.pi, 100)

# unit circle
r = np.sqrt(1.0)
z1 = r*np.cos(alphas)
z2 = r*np.sin(alphas)
Z = np.array([z1, z2]).T # data of unit circle
# scale
S = np.array([[2, 0],
              [0, 0.5]])
thetas = np.array([0, 30, 45, 60, 90, 120])

for theta in thetas:
    # rotate
    print('==== Rotate ====')
    print(theta)
    theta = theta/180*np.pi
    R = np.array([[np.cos(theta), -np.sin(theta)],
                  [np.sin(theta),  np.cos(theta)]])

    # translate
    c = np.array([2, 1])
    X = Z@S@R.T + c;
    Q = R@np.linalg.inv(S)@np.linalg.inv(S)@R.T
    print('==== Q ====')
    print(Q)
    LAMBDA, V = np.linalg.eig(Q)
    print('==== LAMBDA ====')
    print(LAMBDA)
    print('==== V ====')
    print(V)

    x1 = X[:,0]
    x2 = X[:,1]

    fig, ax = plt.subplots(1)
    ax.plot(z1, z2, 'b') # plot the unit circle
    ax.plot(x1, x2, 'r') # plot the transformed shape
    ax.plot(c[0],c[1],'xk') # plot the center

    ax.quiver(0,0,1,0,color = 'b',angles='xy', scale_units='xy',scale=1)
    ax.quiver(0,0,0,1,color = 'b',angles='xy', scale_units='xy',scale=1)
    ax.quiver(0,0,-1,0,color = 'b',angles='xy', scale_units='xy',scale=1)
    ax.quiver(0,0,0,-1,color = 'b',angles='xy', scale_units='xy',scale=1)

    ax.quiver(0,0,c[0],c[1],color = 'k',angles='xy', scale_units='xy',scale=1)

    ax.quiver(c[0],c[1],
              V[0,0]/np.sqrt(LAMBDA[0]),
              V[1,0]/np.sqrt(LAMBDA[0]),color = 'r',
              angles='xy', scale_units='xy',scale=1)

    ax.quiver(c[0],c[1],
              V[0,1]/np.sqrt(LAMBDA[1]),
              V[1,1]/np.sqrt(LAMBDA[1]),color = 'r',
              angles='xy', scale_units='xy',scale=1)

    ax.quiver(c[0],c[1],
              -V[0,0]/np.sqrt(LAMBDA[0]),
              -V[1,0]/np.sqrt(LAMBDA[0]),color = 'r',
              angles='xy', scale_units='xy',scale=1)

    ax.quiver(c[0],c[1],
              -V[0,1]/np.sqrt(LAMBDA[1]),
              -V[1,1]/np.sqrt(LAMBDA[1]),color = 'r',
              angles='xy', scale_units='xy',scale=1)
```

```python
plt.axvline(x=0, color= 'k', zorder=0); plt.axhline(y=0, color= 'k', zorder=0)

ax.set_aspect(1)
plt.xlim(-2,4); plt.ylim(-2,4)
plt.grid(linestyle='--')
plt.show(); plt.xlabel('$x_1$'); plt.ylabel('$x_2$')
```

多元高斯分布

多元高斯分布的概率密度函数解析式如下：

$$f_{\chi}(\boldsymbol{x}) = \frac{\exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right)}{(2\pi)^{\frac{D}{2}}\left|\boldsymbol{\Sigma}\right|^{\frac{1}{2}}} \tag{39}$$

相信大家已经在它的分母中看到了本节给出的旋转椭圆解析式。这就是为什么多元高斯分布和很多基于高斯分布的机器学习算法能够和以椭圆为代表的圆锥曲线扯上关系。

(39) 中 $\boldsymbol{\Sigma}$ 的不同形态还会影响到椭圆的形状，如图 4 所示。这是我们要在本系列丛书《概率哦统计》一册深入探讨的内容。

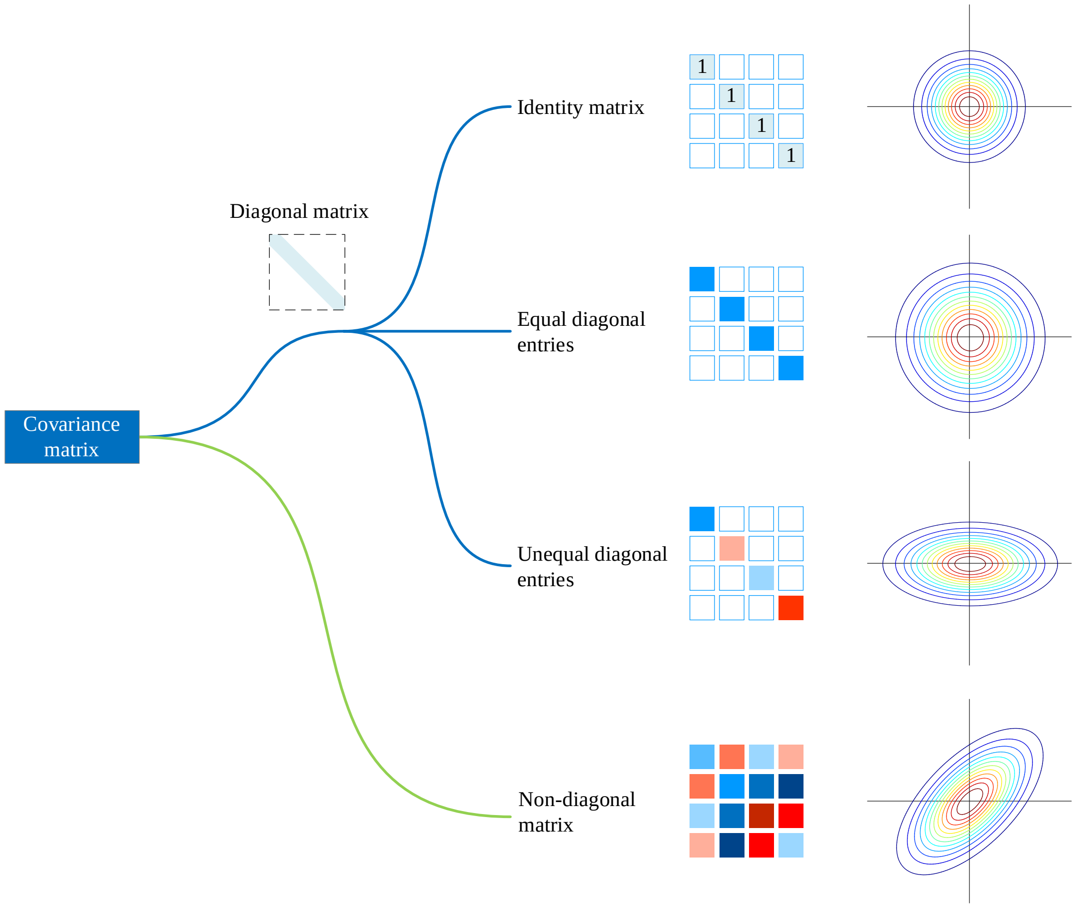


图 4. 协方差矩阵的形态影响高斯密度函数形状

20.4 从单位双曲线到旋转双曲线

循着上一节思路，本节讲解如何从单位双曲线变换得到任意双曲线。平面上，单位双曲线定义如下：

$$\boldsymbol{z}^{\mathrm{T}}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\boldsymbol{z}-1=0 \tag{40}$$

展开得到：

$$z_1^2 - z_2^2 = 1 \tag{41}$$

目标双曲线的解析式为：

$$\frac{(x_1-c_1)^2}{a^2}-\frac{(x_2-c_2)^2}{b^2}=1 \tag{42}$$

和前文思路完全一致，首先对 $\boldsymbol{z}$ 通过 $\boldsymbol{S}$ 缩放，再通过 $\boldsymbol{R}$ 逆时针旋转 θ，最后平移 $\boldsymbol{c}$：

$$\underbrace{\boldsymbol{R}}_{\text{Rotate}}\ \underbrace{\boldsymbol{S}}_{\text{Scale}}\boldsymbol{z}+\underbrace{\boldsymbol{c}}_{\text{Translate}}=\boldsymbol{x} \tag{43}$$

同样展开得到：

$$\underbrace{\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}}_{\text{Rotate}}\underbrace{\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}}_{\text{Scale}}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}+\underbrace{\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}}_{\text{Translate}}=\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{44}$$

后续推导和上一节完全一致，我们可以得到任意双曲线的解析式。

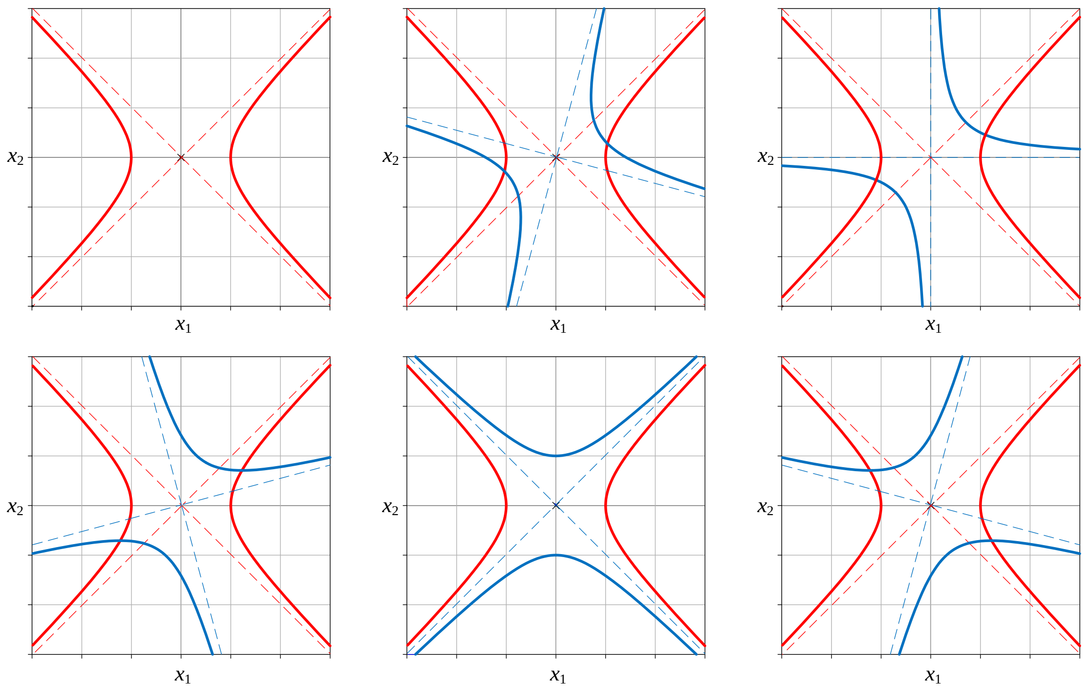


图 5. 通过单位双曲线旋转得到的一系列双曲线

图 5 所示为通过单位双曲线旋转得到的一系列双曲线。注意，图 5 中蓝色双曲线仅仅受到旋转几何操作，没有经过缩放和平移操作。

请大家自行修改 Bk4_Ch20_02.py 参数绘制不同几何变换条件下双曲线。

```
# Bk4_Ch20_02.py

import numpy as np
import matplotlib.pyplot as plt

alphas = np.linspace(0, 2*np.pi, 100)

# unit circle
r = np.sqrt(1.0)

z1 = r*1/np.cos(alphas)
z2 = r*np.tan(alphas)

Z = np.array([z1, z2]).T # data of unit circle

# scale
S = np.array([[1, 0],
              [0, 1]])

thetas = np.array([0, 30, 45, 60, 90, 120])

for theta in thetas:

    # rotate
    print('==== Rotate ====')
    print(theta)
    theta = theta/180*np.pi
    R = np.array([[np.cos(theta), -np.sin(theta)],
                  [np.sin(theta),  np.cos(theta)]])

    X = Z@S@R.T;

    x1 = X[:,0]
    x2 = X[:,1]

    fig, ax = plt.subplots(1)
    ax.plot(z1, z2, 'b') # plot the unit circle
    ax.plot(x1, x2, 'r') # plot the transformed shape

    plt.axvline(x=0, color= 'k', zorder=0)
    plt.axhline(y=0, color= 'k', zorder=0)
    ax.set_aspect(1)
    plt.xlim(-3,3)
    plt.ylim(-3,3)
    plt.grid(linestyle='--')
    plt.show()
    plt.xlabel('$x_1$')
    plt.ylabel('$x_2$')
```

20.5 切线：构造函数，求梯度向量

本节探讨如何求解圆锥曲线切线解析式。

椭圆

首先以椭圆为例求解其切线解析式。标准椭圆解析式为：

$$\frac{x_1^2}{a^2}+\frac{x_2^2}{b^2}=1 \tag{45}$$

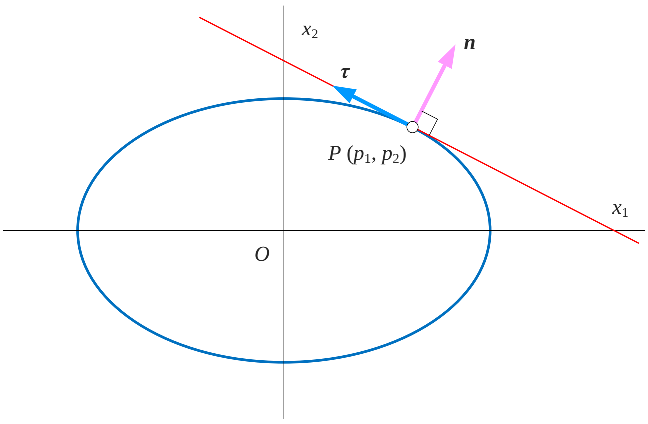


图 6. 椭圆上点 P 处切向量和法向量

先构造一个二元函数 $f(x_1, x_2)$，如下：

$$f\left(x_1,x_2\right)=\frac{x_1^2}{a^2}+\frac{x_2^2}{b^2} \tag{46}$$

如图 6 所示，椭圆上 $P\ (p_1, p_2)$ 一点处 $f(x_1, x_2)$ 梯度，即法向量 $\boldsymbol{n}$ 为：

$$\boldsymbol{n}=\nabla f\left(\boldsymbol{x}\right)\Big|_{(p_1,p_2)}=\left.\begin{bmatrix}\dfrac{\partial f}{\partial x_1}\\ \dfrac{\partial f}{\partial x_2}\end{bmatrix}\right|_{(p_1,p_2)}=\begin{bmatrix}\dfrac{2p_1}{a^2}\\ \dfrac{2p_2}{b^2}\end{bmatrix} \tag{47}$$

如图 6 所示，切线上任意一点和点 P 构成向量，垂直于法向量 $\boldsymbol{n}$，因此两者内积为 0，即，

$$\begin{bmatrix}\dfrac{2p_1}{a^2}\\ \dfrac{2p_2}{b^2}\end{bmatrix}\cdot\begin{bmatrix}x_1-p_1\\ x_2-p_2\end{bmatrix}=\frac{2p_1}{a^2}\left(x_1-p_1\right)+\frac{2p_2}{b^2}\left(x_2-p_2\right)=0 \tag{48}$$

整理上式，得到 $P\ (p_1, p_2)$ 点处椭圆切线解析式：

$$\frac{p_1}{a^2}x_1+\frac{p_2}{b^2}x_2=\frac{p_1^2}{a^2}+\frac{p_2^2}{b^2} \tag{49}$$

图 7 所示为某个给定椭圆上不同点切线。

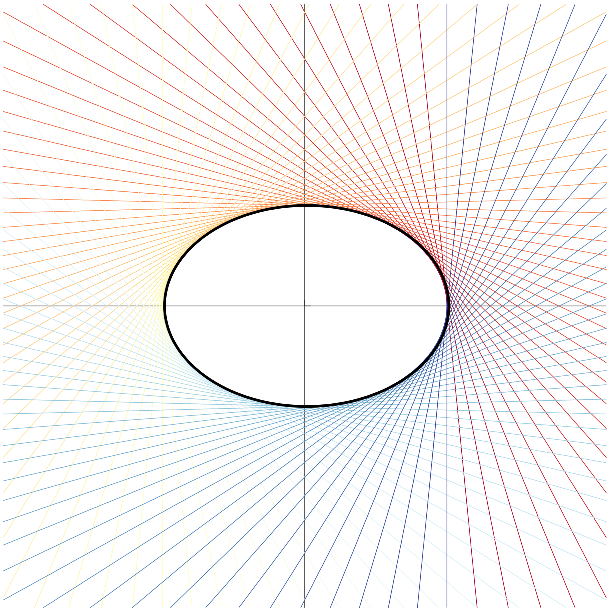


图 7. 椭圆切线分布

正圆

正圆是椭圆特殊形式，将 $a = b = r$ 带入上式，可获得圆心位于原点的正圆上 P (p_1, p_2) 点切线解析式：

$$p_1x_1 + p_2x_2 = p_1^2 + p_2^2 = r^2 \tag{50}$$

图 8 所示为中心位于原点单位圆不同点上切线。

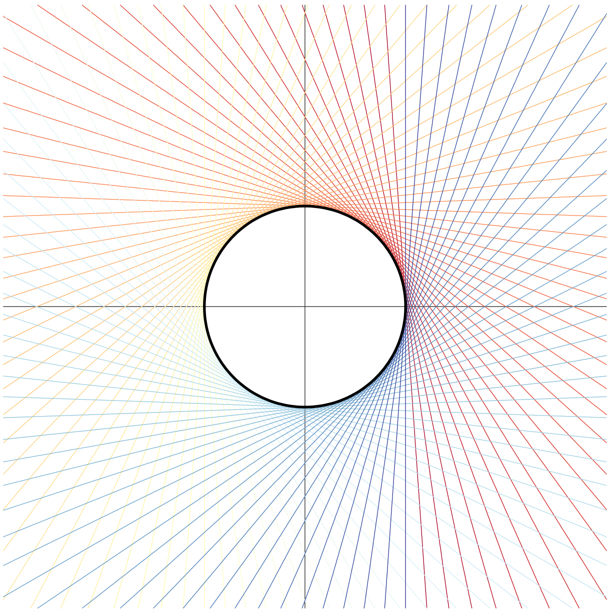


图 8. 单位圆切线分布

双曲线

同样的方法可以求解标准双曲线切线。标准椭圆解析式写作：

$$\frac{x_1^2}{a^2}-\frac{x_2^2}{b^2}=1 \tag{51}$$

类似地，先构造一个二元函数 $f(x_1, x_2)$，如下：

$$f\left(x_1,x_2\right)=\frac{x_1^2}{a^2}-\frac{x_2^2}{b^2} \tag{52}$$

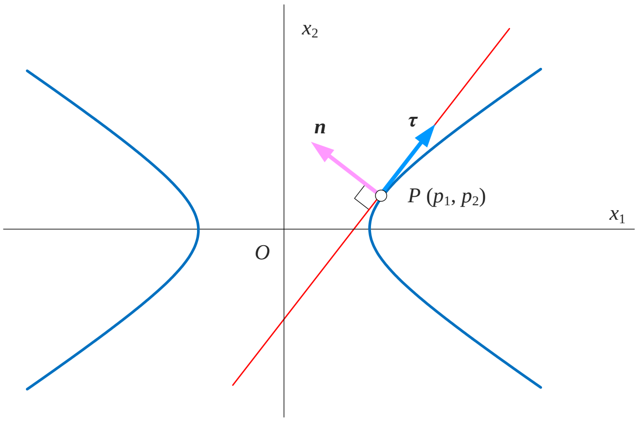


图 9. 双曲线上点 P 处切向量和法向量

如图 9 所示，双曲线上 $P\ (p_1, p_2)$ 点处函数 $f(x_1, x_2)$ 梯度，即法向量 $\boldsymbol{n}$ 为：

$$\boldsymbol{n}=\nabla f\left(\boldsymbol{x}\right)\Big|_{(p_1,p_2)}=\begin{bmatrix}\dfrac{\partial f}{\partial x_1}\\ \dfrac{\partial f}{\partial x_2}\end{bmatrix}\Bigg|_{(p_1,p_2)}=\begin{bmatrix}\dfrac{2p_1}{a^2}\\ -\dfrac{2p_2}{b^2}\end{bmatrix} \tag{53}$$

如图 9 所示，切线上任意一点和点 P 构成向量，垂直于法向量 $\boldsymbol{n}$，通过内积为 0 得到以下等式：

$$\begin{bmatrix}\dfrac{2p_1}{a^2}\\ -\dfrac{2p_2}{b^2}\end{bmatrix}\cdot\begin{bmatrix}x_1-p_1\\ x_2-p_2\end{bmatrix}=\frac{2p_1}{a^2}\left(x_1-p_1\right)-\frac{2p_2}{b^2}\left(x_2-p_2\right)=0 \tag{54}$$

整理上式，得到双曲线上 $P\ (p_1, p_2)$ 点处切线解析式：

$$\frac{p_1}{a^2}x_1-\frac{p_2}{b^2}x_2=\frac{p_1^2}{a^2}-\frac{p_2^2}{b^2} \tag{55}$$

图 10 展示单位双曲线不同点切线。

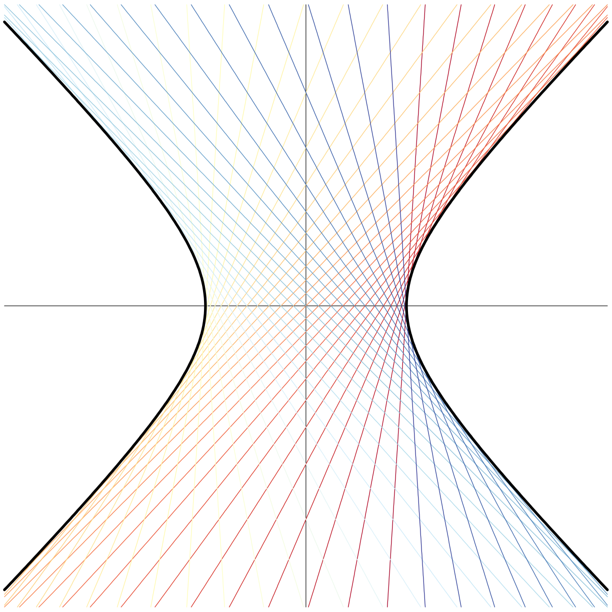


图 10. 双曲线左右两侧切线分布

圆锥曲线一般式

本章前文给出圆锥曲线常见一般表达式，同样据此构造一个二元函数 $f(x_1, x_2)$:

$$f\left(x_1,x_2\right)=Ax_1^2+Bx_1x_2+Cx_2^2+Dx_1+Ex_2+F \tag{56}$$

圆锥曲线任意一点 P (p_1, p_2) 处二元函数 $f(x_1, x_2)$ 梯度，即法向量 $\boldsymbol{n}$ 为:

$$\boldsymbol{n}=\nabla f\left(\boldsymbol{x}\right)\Big|_{\left(p_1,p_2\right)}=\begin{bmatrix}\dfrac{\partial f}{\partial x_1}\\\dfrac{\partial f}{\partial x_2}\end{bmatrix}\Bigg|_{\left(p_1,p_2\right)}=\begin{bmatrix}2Ap_1+Bp_2+D\\Bp_1+2Cp_2+E\end{bmatrix} \tag{57}$$

切线上任意一点和点 P 构成向量，垂直于法向量 $\boldsymbol{n}$，因此两者内积为 0:

$$\begin{bmatrix}2Ap_1+Bp_2+D\\Bp_1+2Cp_2+E\end{bmatrix}\cdot\begin{bmatrix}x_1-p_1\\x_2-p_2\end{bmatrix}=0 \tag{58}$$

即

$$\left(2Ap_1+Bp_2+D\right)\left(x_1-p_1\right)+\left(Bp_1+2Cp_2+E\right)\left(x_2-p_2\right)=0 \tag{59}$$

整理得到圆锥曲线任意一点 P (p_1, p_2) 处切线解析式:

$$\left(2Ap_1+Bp_2+D\right)x_1+\left(Bp_1+2Cp_2+E\right)x_2=2Ap_1^2+2Bp_1p_2+2Cp_2^2+Dp_1+Ep_2 \tag{60}$$

Bk4_Ch20_03.py 绘制图 7，请大家修改代码自行绘制本节和下一节其他图像。

```
# Bk4_Ch20_03.py

import numpy as np
import matplotlib.pyplot as plt

a = 1.5
b = 1

x1 = np.linspace(-3,3,200)
x2 = np.linspace(-3,3,200)
xx1,xx2 = np.meshgrid(x1,x2)

fig, ax = plt.subplots()

theta_array = np.linspace(0,2*np.pi,100)

plt.plot(a*np.cos(b*np.sin(theta)),b*np.sin(b*np.sin(theta)),color = 'k')

colors = plt.cm.RdYlBu(np.linspace(0,1,len(theta_array)))

for i in range(len(theta_array)):

    theta = theta_array[i]

    p1 = a*np.cos(theta)
    p2 = b*np.sin(theta)

    tangent = p1*xx1/a**2 + p2*xx2/b**2 - p1**2/a**2 - p2**2/b**2

    colors_i = colors[int(i),:]

    ax.contour(xx1,xx2,tangent, levels = [0], colors = [colors_i])

plt.axis('scaled')
ax.set_xlim(-3,3)
ax.set_ylim(-3,3)
ax.spines['top'].set_visible(False)
ax.spines['right'].set_visible(False)
ax.spines['bottom'].set_visible(False)
ax.spines['left'].set_visible(False)
ax.axvline(x=0,color = 'k')
ax.axhline(y=0,color = 'k')
```

20.6 法线：法向量垂直于切向量

椭圆

(46) 所示标准椭圆上 P (p_1, p_2) 一点处法向量 $\boldsymbol{\tau}$ 为：

$$\boldsymbol{\tau} = \left. \begin{bmatrix} \dfrac{\partial f}{\partial x_2} \\ -\dfrac{\partial f}{\partial x_1} \end{bmatrix} \right|_{(p_1, p_2)} = \begin{bmatrix} \dfrac{2p_2}{b^2} \\ -\dfrac{2p_1}{a^2} \end{bmatrix} \tag{61}$$

$P\ (p_1, p_2)$ 处法向量 $\boldsymbol{\tau}$ 显然垂直其切向量 $\boldsymbol{n}$。

容易获得 $P\ (p_1, p_2)$ 点，椭圆法线解析式：

$$\begin{bmatrix} \dfrac{2p_2}{b^2} \\ -\dfrac{2p_1}{a^2} \end{bmatrix} \cdot \begin{bmatrix} x_1 - p_1 \\ x_2 - p_2 \end{bmatrix} = \frac{2p_2}{b^2}(x_1 - p_1) - \frac{2p_1}{a^2}(x_2 - p_2) = 0 \tag{62}$$

整理得到：

$$\frac{p_2}{b^2}x_1 - \frac{p_1}{a^2}x_2 = \frac{p_1 p_2}{b^2} - \frac{p_1 p_2}{a^2} \tag{63}$$

图 11 所示为椭圆法线分布情况。

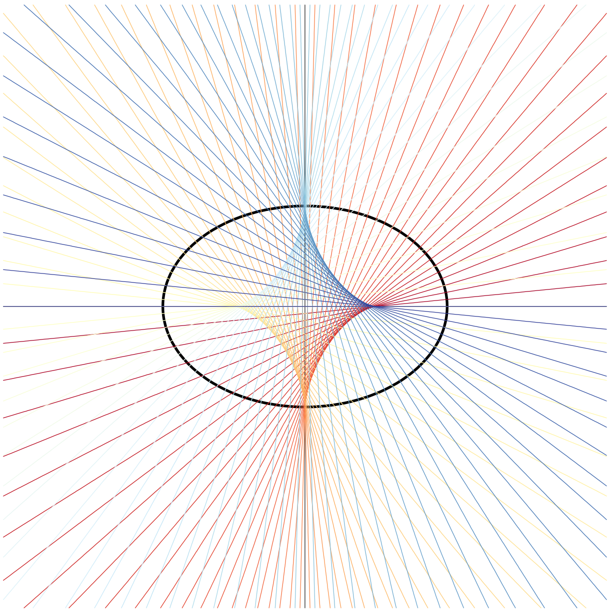


图 11. 椭圆法线分布

圆锥曲线一般式

下面推导一般圆锥曲线的法线。(56) 圆锥曲线解析式上 P 点处切向量 $\boldsymbol{\tau}$ 为：

$$\boldsymbol{\tau} = \begin{bmatrix} \dfrac{\partial f}{\partial x_2} \\ -\dfrac{\partial f}{\partial x_1} \end{bmatrix}\Bigg|_{(p_1,p_2)} = \begin{bmatrix} Bp_1 + 2Cp_2 + E \\ -(2Ap_1 + Bp_2 + D) \end{bmatrix} \tag{64}$$

得到过 P 点圆锥曲线法线直线方程，如下：

$$\begin{bmatrix} Bp_1 + 2Cp_2 + E \\ -(2Ap_1 + Bp_2 + D) \end{bmatrix} \cdot \begin{bmatrix} x_1 - p_1 \\ x_2 - p_2 \end{bmatrix} = 0 \tag{65}$$

整理得到如下法线解析式：

$$(Bp_1 + 2Cp_2 + E)x_1 - (2Ap_1 + Bp_2 + D)x_2 = B(p_1^2 - p_2^2) + (2C - 2A)p_1 p_2 + Ep_1 - Dp_2 \tag{66}$$

下面这幅图最能总结本章的核心内容。它虽然是四副子图，却代表着一个连贯的几何变换操作。不管是从旋转椭圆到单位圆，还是从单位圆到旋转椭圆，请大家务必记住每步几何变换对应的线性代数运算。

理解这些几何变换对于理解协方差矩阵、多元高斯分布、主成分分析和很多机器学习算法有至关重要的作用。

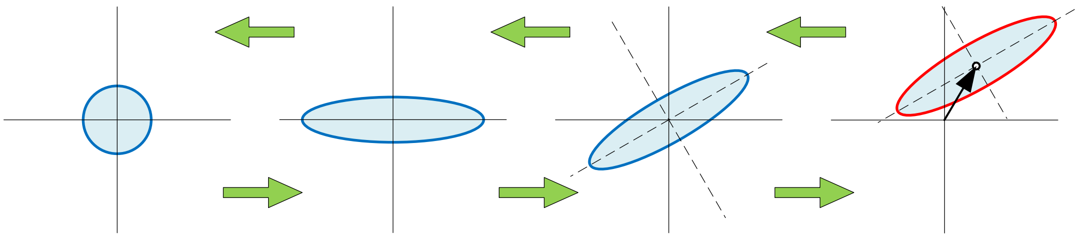


图 12. 总结本章重要内容的四副图