矩阵二次型

Dezeming Family

2021 年 5 月 15 日

DezemingFamily 系列书和小册子因为是电子书，所以可以很方便地进行修改和重新发布。如果您获得了 DezemingFamily 的系列书，可以从我们的网站 [https://dezeming.top/] 找到最新版。对书的内容建议和出现的错误欢迎在网站留言。

目录

一　二次型的概念

首先，假如我们有 n 个变量 x_i 构成一个多项式，其中每一项最高次幂为 2，可以表示如下：

$$f(x_1,x_2,\dots,x_n)=a_{1,1}x_1^2+2a_{1,2}x_1x_2+\dots+a_{n,n}x_n^2 \tag{一.1}$$

如果二次型多项式中只有某变量的平方项，称为二次型的标准型：

$$f(x_1,x_2,\dots,x_n)=a_{1,1}x_1^2+a_{2,2}x_2^2+\dots+a_{n,n}x_n^2 \tag{一.2}$$

二次型多项式可以用矩阵相乘来实现：

$$f(x_1,x_2,\dots,x_n)=\begin{bmatrix}x_1 & x_2 & \dots & x_n\end{bmatrix}\begin{bmatrix}a_{1,1} & a_{1,2} & \dots & a_{1,n}\\ a_{2,1} & a_{2,2} & \dots & a_{2,n}\\ \dots & & & \\ a_{n,1} & a_{n,2} & \dots & a_{n,n}\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ \dots\\ x_n\end{bmatrix} \tag{一.3}$$

$$=X^TAX \tag{一.4}$$

由二次型和二次多项式的基本关系，我们很容易就能发现在二次型矩阵中，$a_{i,j}=a_{j,i}$。

在几何中，在平面直角坐标系下，以坐标原点为圆心的二次曲线的方程一般是下面的形式：

$$a_1x^2+2a_2xy+a_3y^2=b \tag{一.5}$$

所以我们完全可以将矩阵的二次型与二次曲线挂钩。当然，自变量如果只有 3 个以内我们还是可以画出来的，但超过三个就没法直接画出来了。

二　二次型的变换

可逆线性变换

假如对 X ，有可逆矩阵 C 使得 $X=CY$，Y 是向量，对二次型做可逆线性变换：

$$f=X^TAX=(CY)^TA(CY)=Y^T(C^TAC)Y \tag{二.1}$$

因为 $A^T = A$，因此 $(C^TAC)^T = C^TAC$ 也是二次型矩阵。因为 C 可逆，所以 A 与 C^TAC 的秩是相同的。

二次型矩阵的标准型

由矩阵的相似对角化可知，实对称矩阵 A 可以通过正交矩阵 Q 化为对角矩阵：

$$Q^{-1}AQ = Q^TAQ \tag{二.2}$$

因此二次型矩阵也可以这么化为标准型：

$$f = X^TAX = Y^T(Q^TAQ)Y = Y^T\varLambda Y \tag{二.3}$$

$$\varLambda = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} \tag{二.4}$$

其中 QY 为正交变换，λ_i 为矩阵的特征值。

正交变换与正交矩阵

这里简单介绍一下正交变换和正交矩阵。当矩阵 A 满足 $AA^T = A^TA = E$ 时，称 A 为正交矩阵。

对于正交矩阵，我们可以看出（det 表示求行列式）：

$$A^T = A^{-1} \tag{二.5}$$

$$det(A) = 1 \tag{二.6}$$

$$\tag{二.7}$$

且 A 各行（列）之间两两正交，且都是单位向量。正交变换不改变向量的长度，也不改变两个向量的内积：

$$|Ax| = \sqrt{(Ax)^TAx} = \sqrt{x^TA^TAx} = \sqrt{x^Tx} = |x| \tag{二.8}$$

$$\big\langle Ax, Ay \big\rangle = (Ax)^TAy = x^TA^TAy = x^Ty \tag{二.9}$$

而向量的内积其实就是向量相乘再乘以它们之间夹角的 cos 值，因此可以说明两个向量正交变换以后的夹角也不变，因此正交变换不会改变形状。

二次型与椭圆

我们考虑一下该二次型：

$$f(x,y) = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} \frac{5}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{3}{2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{二.10}$$

$$= \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{二.11}$$

我们可以从正交矩阵性质得到：

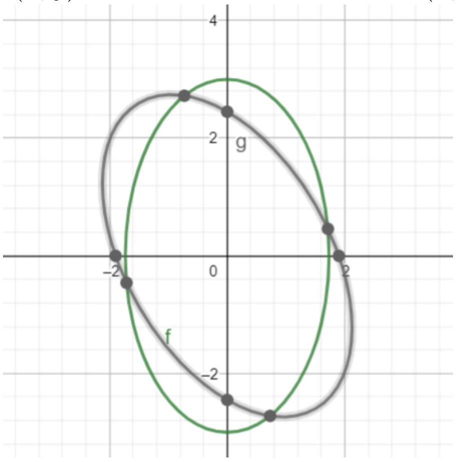
$$\left(\begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} \right) = \left(\begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \right)^T \tag{二.12}$$

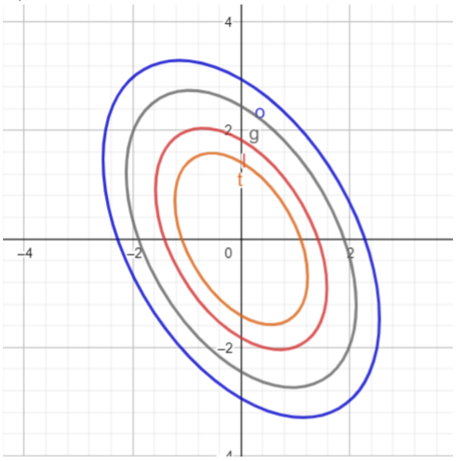
我们设 $f(x,y)$ 为一个定值，比如 $f(x,y) = 9$，得到与 (x,y) 的相关图像显示为灰色；同时我们假设有另一个以 $[x\ y]Q$ 为基底的椭圆 g(u,v)，其中：

$$f(x,y) = 2.5x^2 + \frac{3}{2}y^2 + \sqrt{3}xy = 9 \tag{二.13}$$

$$g(u,v) = 3u^2 + v^2 = 9 \tag{二.14}$$

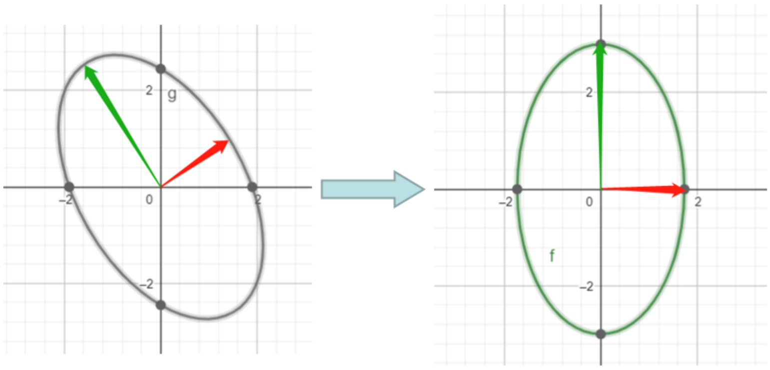
得到的图像显示如下图左，为了方便比较，我们把 f 和 g 显示在同一个图上，注意灰色椭圆横纵坐标表示 (x, y)，绿色椭圆横纵坐标表示为 (u, v)：





神奇（或者说理所当然）的事情发生了，这两个椭圆一模一样！f 椭圆是原本的样子，但当我们对坐标做了正交变换以后，就变成了 g 的样子。

现在看上图的右图，椭圆从小到大，$f(x, y)$ 的值分别为 3、5、9 和 13。我们知道 $f(x, y)$ 是 (x, y) 的函数，我们根据下面示意图可知，当我们给 f(x,y) 设置一个目标值时，从椭圆中心出发的向量 (x, y) 沿着红色箭头位置可以最快达到目标值，而沿着绿色箭头则会最慢达到目标值：



而我们对红色和绿色箭头做了正交变换以后，就相当于我们可以知道，现在从椭圆中心的向量沿着坐标轴可以分别最快或最慢地达到目标值。

注意在某个目标值下的椭圆的长轴和短轴的长度就可以由二次型标准型的特征向量求出，而二次型一般型我们难以直接看出长轴和短轴的长短。

参考文献

[1] Strang G . Introduction to Linear Algebra[M]. Wellesley-Cambridge Press, 2003.