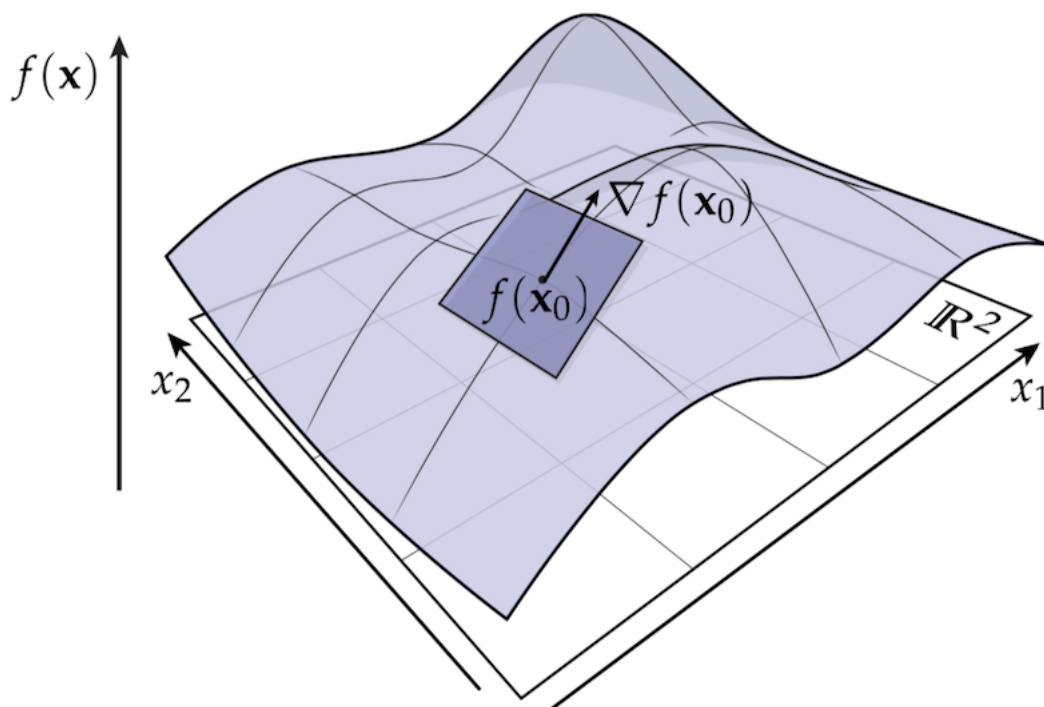


# 梯度、散度、旋度

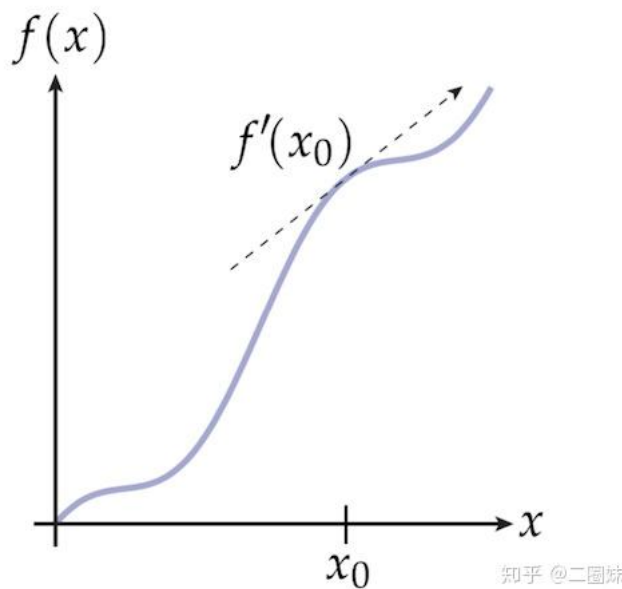
---



## 梯度

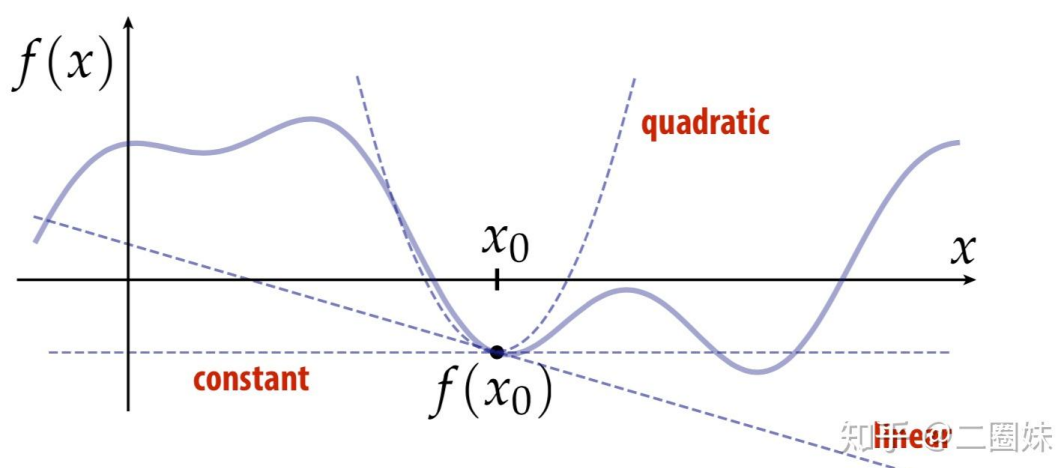
---

如果我们将导数理解为微小变化引起的变化：



泰勒级数如下：

$$f(x) = \overset{\text{constant}}{f(x_0)} + \overset{\text{linear}}{f'(x_0)(x - x_0)} + \overset{\text{quadratic}}{\frac{(x-x_0)^2}{2!}f''(x_0)} + \cdots$$

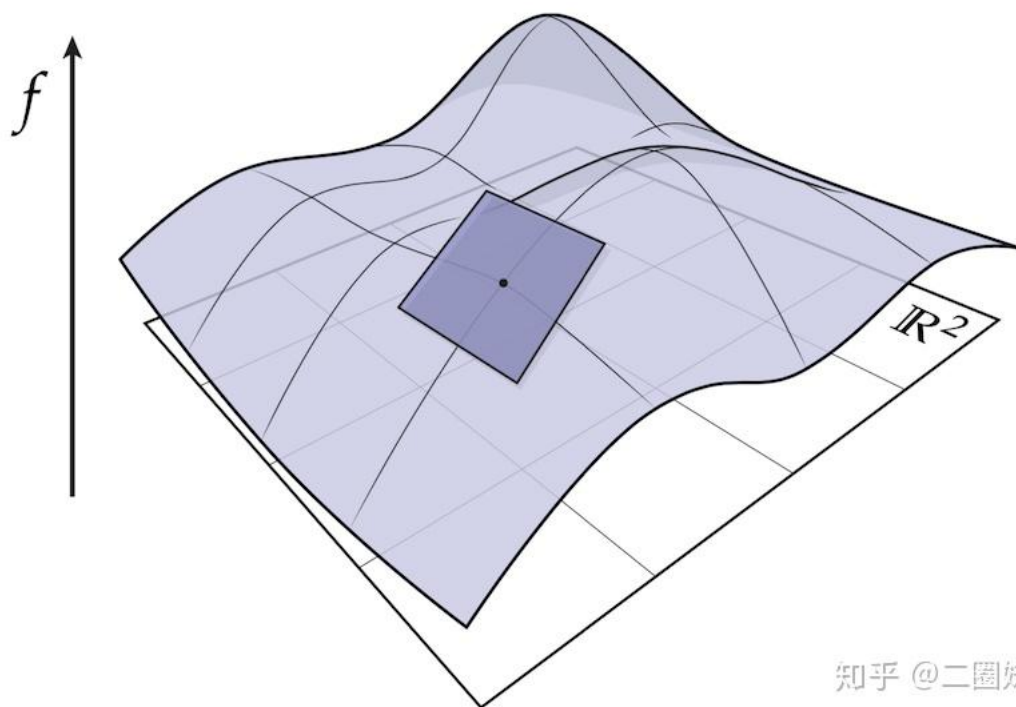


经常因为计算的原因我们可以利用线性的部分来做近似：

$$L(x) = f(x_0) + f'(x_0)(x - x_0)$$

这也 make sense，相当于我们找到  $x_0$  处的切线，然后来看变化，毕竟这就是导数的本质定义。

这种思路对于具有多变量的函数应当也是适用的，比如空间中的曲面  $f(x,y)$ ：



对于这样的曲面，我们如果想要在  $(x_0, y_0)$  附近做线性的近似，那当然应该找到它对应的切平面。而切平面的方程为：

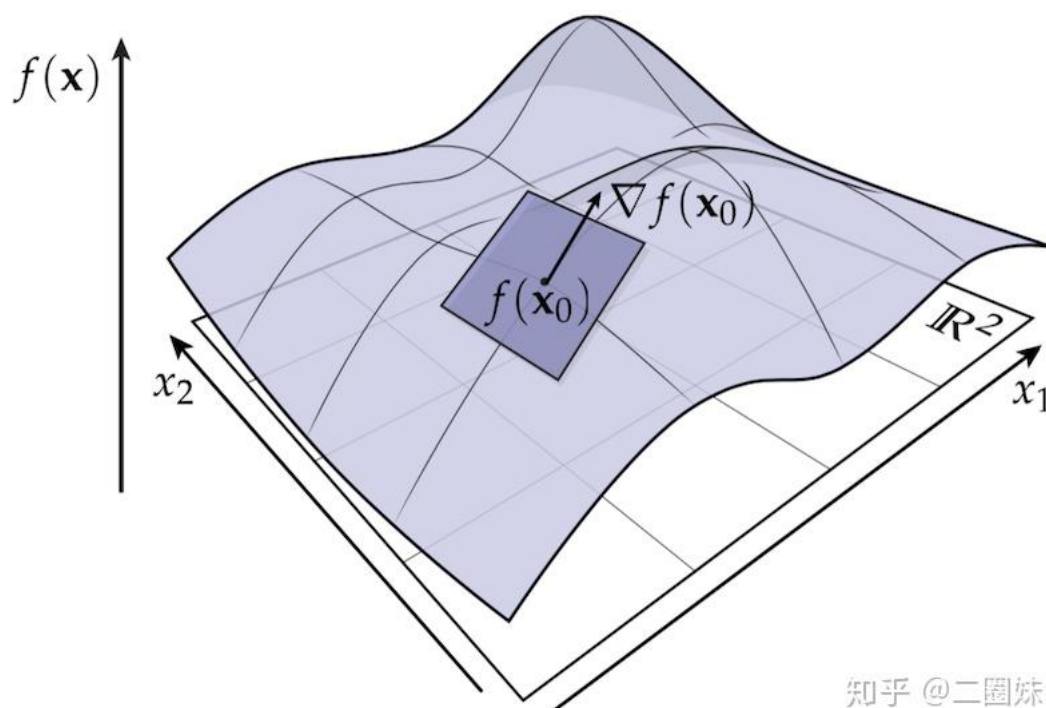
$$L(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$$

（切平面的方程可以通过找到x方向与y方向的切线，然后cross product得到法向量来证明。）

非常重要的一点是，上面这个式子同时也某种程度上回应了偏导数的本质，就是我们看x变量（方向）上的微小变化引起的变化，再加上y变量（方向）的微小变化引起的变化，就是总的变化。即使维度增加，线性模拟的式子应该也是类似的，比如函数  $f(x, y, z)$  在  $(a, b, c)$  处的线性近似应该是：

$$L(x, y, z) = f(a, b, c) + f_x(a, b, c)(x - a) + f_y(a, b, c)(y - b) + f_z(a, b, c)(z - c)$$

这也帮我们自然而然的引出了**梯度 (gradient)** 的概念：



知乎 @二圈妹

线性近似写成向量形式：

$$L(\mathbf{x}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)$$

梯度就定义为：

$$\mathbf{grad} f = \frac{\partial f}{\partial x} \mathbf{i} + \frac{\partial f}{\partial y} \mathbf{j}$$

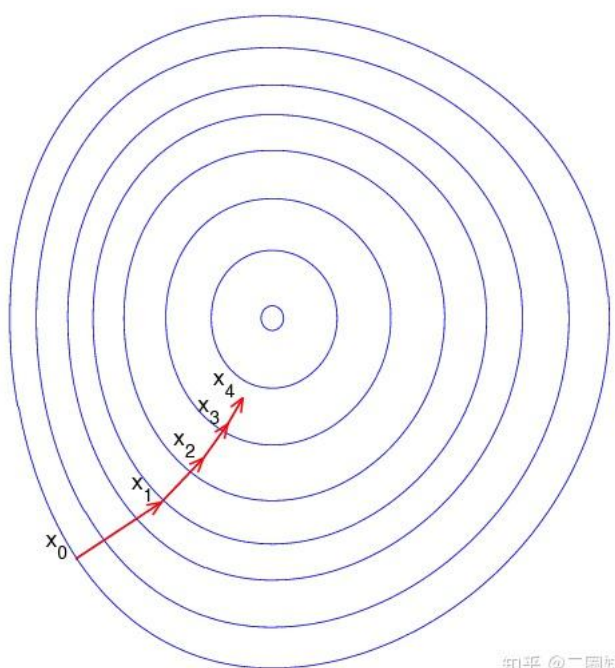
Nabla算子：

$$\nabla = \frac{\partial}{\partial x} \mathbf{i} + \frac{\partial}{\partial y} \mathbf{j}$$

以上式子也告诉我们，如果我们从  $\mathbf{x}_0$  处动身：

- 沿着梯度方向走，函数值增大
- 沿着相反于梯度的方向走，函数值减小
- 垂直于梯度方向，函数值不变

同时这也呼应了许多算法，比如‘梯度下降法’，因为朝着相反梯度的方向，就是下降的最大的方向，同时因为以上我们指出的‘垂直于梯度方向，函数值不变’，所以我们如果在等高线上看，我们总是朝着与等高线垂直的方向走：



当然多元函数可以更多元，比如对于3d空间中的温度函数  $f(x,y,z)$ ，我们也可以求它的梯度，它的梯度我们计算出来就是一个三维空间中的向量场，这也是一个函数，如果我们代入空间中的每一点的数值，那么就可以知道其变化最快的方向和大小。当然也可以继续更多维度的函数  $f(x_1, x_2, \dots, x_n)$ 。

给我们一个多元函数，梯度的作用就是给了表示它变化的向量场。所以我觉得梯度跟导数更像一家人。

继续来看二次近似：

$$Q(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2$$

如果我们想把它推广到多元函数，那么我们会想应该也是类似的状况，除了前面的这一部分  $L(x, y) = a + b(x - x_0) + c(y - y_0)$ ，我们会想要再加上二次的部分： $d(x - x_0)^2 + e(x - x_0)(y - y_0) + f(y - y_0)^2$ 。

$$Q(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) + \frac{1}{2}f_{xx}(x_0, y_0)(x - x_0)^2 + f_{xy}(x_0, y_0)(x - x_0)(y - y_0) + \frac{1}{2}f_{yy}(x_0, y_0)(y - y_0)^2$$

通过类比或者计算，我们都可以得到以上的二次近似的结果。

然后我们可以继续引出二次近似的向量形式：

$$Q(\mathbf{x}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^T H(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)$$

其中 H 是 Hessian 矩阵：

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}$$

当然我们也可以推广到更高的维度：

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

Hessian 矩阵必定是一个对称矩阵（比如之前的线性近似这是内积

$\nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)$ ，按照内积的定义其实是看两个向量的对齐程度，这里我们某种程度上做的也是这件事，而在代表内积时，矩阵总是对称的，所以这里的矩阵也必定是对称的）。

来看一下关于内积的梯度：

$$f = \mathbf{u}^T \mathbf{v}$$

如果这样来看：

$$\mathbf{u}^T \mathbf{v} = \sum_{i=1}^n u_i v_i$$

$$\frac{\partial}{\partial u_k} \sum_{i=1}^n u_i v_i = \sum_{i=1}^n \frac{\partial}{\partial u_k} (u_i v_i) = v_k$$

所以这个结果就是：

$$\nabla_{\mathbf{u}}(\mathbf{u}^T \mathbf{v}) = \mathbf{v}$$

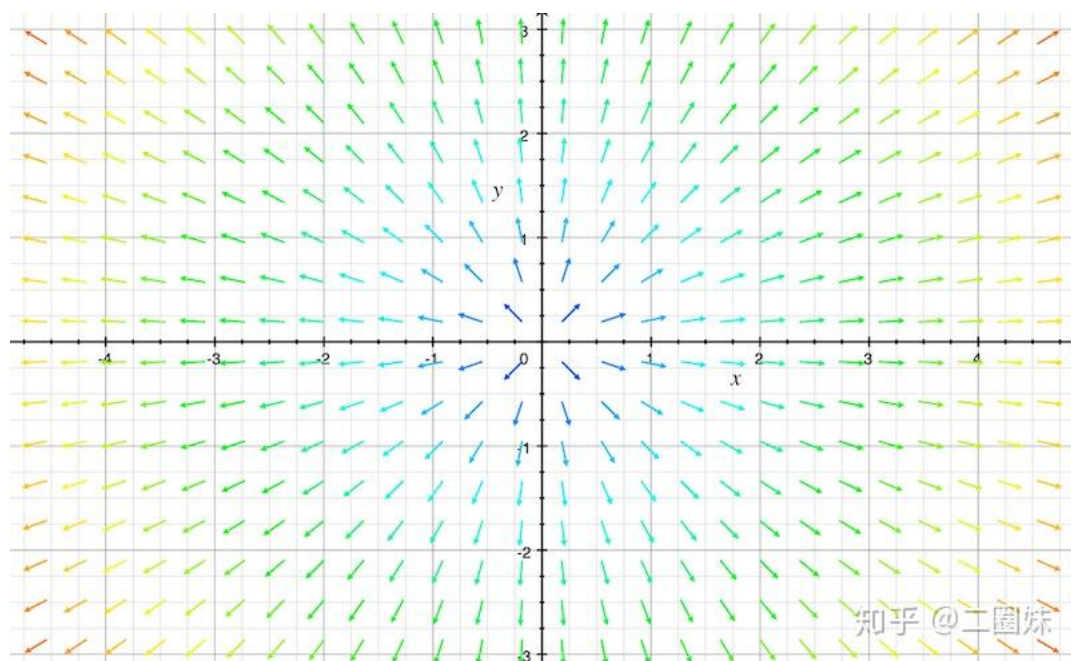
其实跟就跟  $\frac{d}{dx}(xy) = y$  差不多。

所以有一些对矩阵/向量求梯度的公式，都能找到对应的亲戚：

矩阵求导	亲戚
$\nabla_{\mathbf{x}}(\mathbf{x}^T \mathbf{y}) = \mathbf{y}$	$\frac{d}{dx}(xy) = y$
$\nabla_{\mathbf{x}}(\mathbf{x}^T \mathbf{x}) = 2\mathbf{x}$	$\frac{d}{dx}(x^2) = 2x$
$\nabla_{\mathbf{x}}(\mathbf{x}^T A \mathbf{y}) = A \mathbf{y}$	$\frac{d}{dx}(axy) = ay$
$\nabla_{\mathbf{x}}(\mathbf{x}^T A \mathbf{x}) = 2A \mathbf{x}$	$\frac{d}{dx}(ax^2) = 2ax$
...	...

同时这也在呼应之前的说法‘梯度和导数更像一家人’。o(∩\_∩)o

多元函数  $f(x_1, \dots, x_n)$  带来了梯度，梯度的表现方式是向量场，这个场告诉我们的的是空间中的每一个点这个多元函数变化的大小和方向。



画向量场，一般用箭头表示它的指向，箭头长度表示它的大小，但是因为如果画出实际的大小可能会很乱，所以有些是用颜色来代表大小，越暖（红）就是此处的向量越大，越冷（蓝）就是越小。

向量场当然也是函数，因为对于空间中的每一点，都有一个对应的箭头，比如上图：



$$X : R^2 \rightarrow R^2$$

实际生活中有许多向量场：电场、磁场、重力场、速度场。

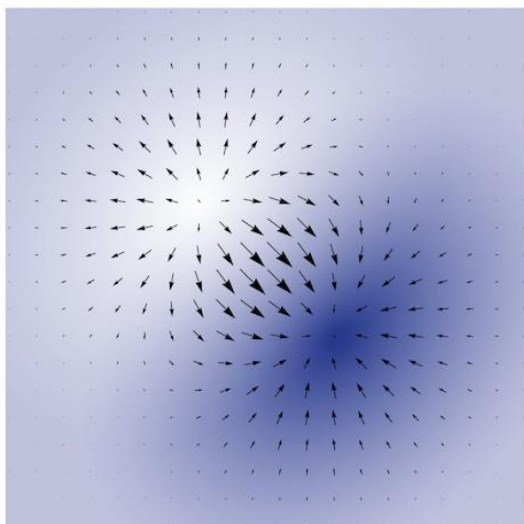
我喜欢将向量场想象为流动着的液体，这些液体在某一点的流速就是对应的箭头，这些液体当然不一定需要遵守物理规律，它们可以凭空出现或者消失，也可以随意的变换运动的方向。

那么这样会有两个相关的问题产生：

- 这些液体只是在空间中运动，空间中某点会流出液体么？（就像喷泉那样🌊）或者这些液体会流入（像地上有个洞，液体会从这里流走）？
- 这些液体会在空间中旋转么？就像所谓的湍流或者龙卷风那样？

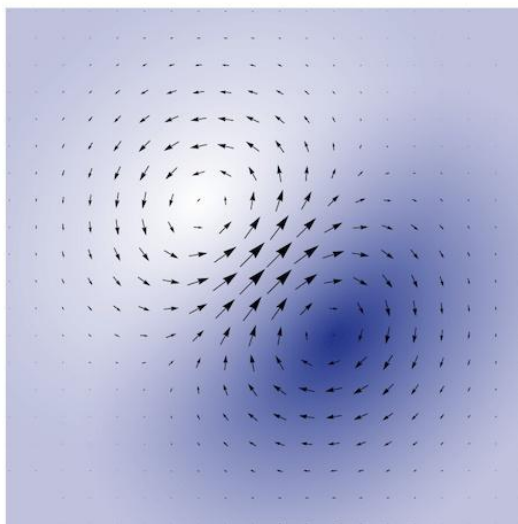
这两个问题我们都可以通过分析向量场而知道，从而引出了散度和旋度的概念：

**“How much is field shrinking/expanding?”**



$\text{div } X$

**“How much is field spinning?”**

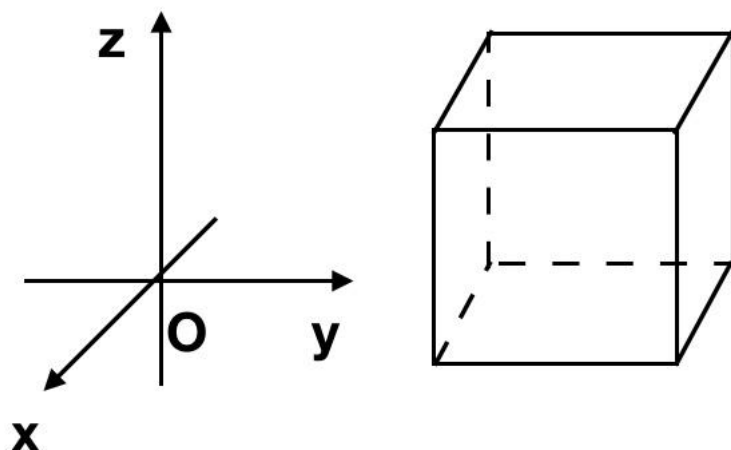


$\text{curl } Y$  知乎 @二圈妹

## 散度

像上面提到的，散度这个是用来表明空间中的这个点是否产生液体，其实也就是来看在此点处向内的箭头比较多还是向外的箭头比较多，我们现在来推导散度的表达式。

假设空间中有向量场  $\mathbf{F} = F_x \mathbf{i} + F_y \mathbf{j} + F_z \mathbf{k}$ ，那么观察空间中的微小的长方体：



知乎 @二圈妹

我们先就处理  $x$  轴方向的出入箭头之差,先看正面, 如果这个长方体极小的话, 我们可以做以下近似:

$$\iint_{\text{frontside}} F_x(x, y, z) dS \simeq F_x\left(x + \frac{\Delta x}{2}, y, z\right) \Delta y \Delta z$$

$$\iint_{\text{backside}} F_x(x, y, z) dS \simeq -F_x\left(x - \frac{\Delta x}{2}, y, z\right) \Delta y \Delta z$$

所以

$$\begin{aligned} \iint_{\text{front+back}} F_x(x, y, z) dS &= F_x\left(x + \frac{\Delta x}{2}, y, z\right) \Delta y \Delta z - F_x\left(x - \frac{\Delta x}{2}, y, z\right) \Delta y \Delta z \\ &= \frac{F_x\left(x + \frac{\Delta x}{2}, y, z\right) \Delta y \Delta z - F_x\left(x - \frac{\Delta x}{2}, y, z\right) \Delta y \Delta z}{\Delta x} \Delta x \end{aligned}$$

这是针对长方体的, 那么如果我们要长方体中的一点的话就是来看它的密度

$\Delta V = \Delta x \Delta y \Delta z$ :

$$\begin{aligned} \lim_{\Delta V \rightarrow 0} \frac{1}{\Delta V} \iint_{\text{front+back}} F_x(x, y, z) dS &= \lim_{\Delta x \rightarrow 0} \frac{F_x\left(x + \frac{\Delta x}{2}, y, z\right) \Delta y \Delta z - F_x\left(x - \frac{\Delta x}{2}, y, z\right) \Delta y \Delta z}{\Delta x} \\ &= \frac{\partial F_x}{\partial x} \end{aligned}$$

同样, 我们可以推导出  $y$  方向,  $z$  方向是也是类似的, 这也就是我们的散度:

$$\text{div } \mathbf{F} = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z}$$



这就是这个点是否往外产生液体，散度我觉得是很好的翻译，而散度的英文 divergence 也有‘脱离,逸出’的意思，也是很表达它的意思了。

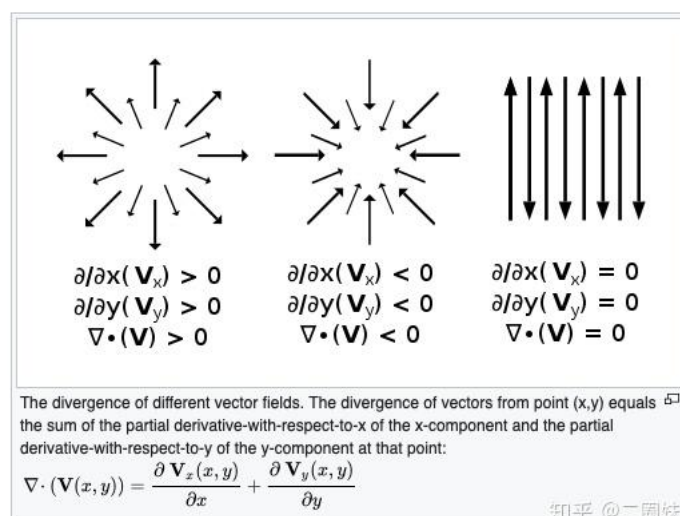
散度是给我们一个向量场，我们得到的是一个标量，代表的是一点是否有‘逸出’。

另外一个理解是既然我们想要知道的是三个方向的向量场的变化，我们当然就会想分别对它们求导，然后看最终的变化，这个散度我们也经常写成：

$$\text{div } \mathbf{F} = \nabla \cdot \mathbf{F}$$

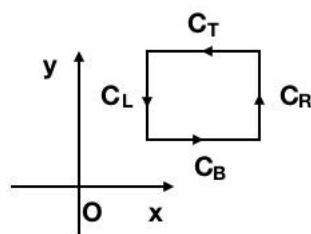
因为把  $\nabla$  算子看做向量运算符，函数本身也是向量，所以是点乘的关系。

wikipedia 上关于散度的图会给我们更加直观的理解：



## 旋度

现在我们来解决第二个问题。考量与  $xy$  平面 平行的一个长方形：



知乎 @二图妹

那么对于旋转的贡献：

$$\begin{aligned}
\int_{C_B} F_x dx &\simeq F_x(x, y - \frac{\Delta y}{2}, z) \Delta x \\
\int_{C_T} F_x dx &\simeq -F_x(x, y + \frac{\Delta y}{2}, z) \Delta x \\
\int_{C_B+C_T} F_x dx &\simeq - \left[ F_x(x, y + \frac{\Delta y}{2}, z) - F_x(x, y - \frac{\Delta y}{2}, z) \right] \Delta x \quad (5) \\
&= - \frac{F_x(x, y + \frac{\Delta y}{2}, z) - F_x(x, y - \frac{\Delta y}{2}, z)}{\Delta y} \Delta x \Delta y \quad (6)
\end{aligned}$$

除以面积，毕竟我们考察的是某点对于旋转的贡献：

$$\frac{1}{\Delta S} \int_{C_B+C_T} F_x dx = - \frac{F_x(x, y + \frac{\Delta y}{2}, z) - F_x(x, y - \frac{\Delta y}{2}, z)}{\Delta y}$$

同理我们看左边和右边：

$$\frac{1}{\Delta S} \int_{C_L+C_R} F_y dy = \frac{F_y(x + \frac{\Delta x}{2}, y, z) - F_y(x - \frac{\Delta x}{2}, y, z)}{\Delta x}$$

所以有：

$$\lim_{\Delta S \rightarrow 0} \frac{1}{\Delta S} = \oint \mathbf{F} \cdot \mathbf{t} ds = \frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}$$

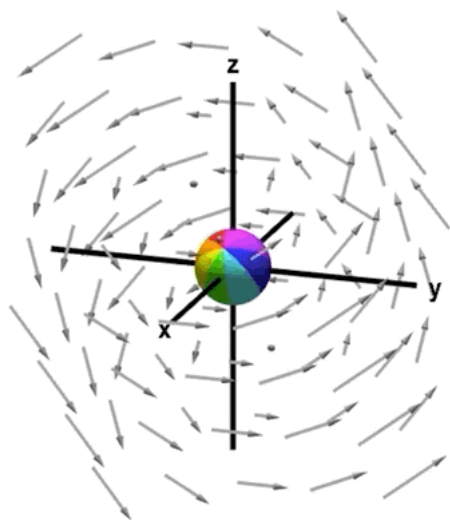
这是平行于xy平面的部分，同理我们可以继续求出平行于xz和yz的部分，最终会得到我们的旋度：

$$\mathbf{curl} \mathbf{F} = \mathbf{i} \left( \frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z} \right) + \mathbf{j} \left( \frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x} \right) + \mathbf{k} \left( \frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right)$$

不是很容易记住，所以这个时候我们就可以感谢  $\nabla$  算子了：

$$\mathbf{curl} \mathbf{F} = \nabla \times \mathbf{F}$$

旋度是给我们一个向量场，我们得到的是一个向量，代表的是这一点的旋转情况。



参考:

- [Vector Calculus](#)