

从零开始一起学习SLAM | 为什么要用齐次坐标?

在涉及到计算机视觉的几何问题中，我们经常看到齐次坐标这个术语。本文介绍一下究竟为什么要用齐次坐标？使用齐次坐标到底有什么好处？

什么是齐次坐标？

简单的说：齐次坐标就是在原有坐标加上一个维度：

$$\begin{aligned}(x, y) &\rightarrow (x, y, 1) \\ (x, y, z) &\rightarrow (x, y, z, 1)\end{aligned}$$



使用齐次坐标有什么优势？

齐次坐标的使用能够大大简化在三维空间中的点线面表达方式和旋转平移等操作，具体分如下几点进行说明。

1 能否非常方便的表达点在直线或平面上

在2D平面上，一条直线 l 可以用方程 $ax + by + c = 0$ 来表示，该直线用向量表示的话一般记做

$$l = (a, b, c)^T$$

我们知道点 $p = (x, y)$ 在直线 l 上的充分必要条件是 $ax + by + c = 0$

如果使用齐次坐标的话，点 p 的齐次坐标就是

$$p' = (x, y, 1)$$

那么 $ax + by + c = 0$ 就可以用两个向量的内积（点乘）来表示：

$$ax + by + c * 1 = (a, b, c)^T (x, y, 1) = l^T * p' = 0$$

因此，点 p 在直线 l 上的充分必要条件就是 直线 l 与 p 的齐次坐标 p' 的内积：

$$l^T * p' = 0$$

是不是很方便呢！

同理，我们知道 三维空间的一个平面 A 可以用方程 $ax + by + cz + d = 0$ 来表示，三维空间的一个点 $P = (x, y, z)$ 的齐次坐标 $P' = (x, y, z, 1)$ ，类似的，点 P 在空间平面 A 上可以用两个向量的内积来表示，如下：

$$\begin{aligned} ax + by + c * z + d &= \\ (a, b, c, d)^T (x, y, z, 1) &= \\ A^T * P' &= 0 \end{aligned}$$

因此，点 P 在平面 A 上的充分必要条件就是平面 A 向量与 P 的齐次坐标 P' 的内积（点乘）：

$$A^T * P' = 0$$

2 方便表达直线与直线，平面与平面的交点

先给出结论，后面再具体解释：

结论：在齐次坐标下，可以用两个点 p, q 的齐次坐标叉乘结果来表达一条直线 l ，也就是

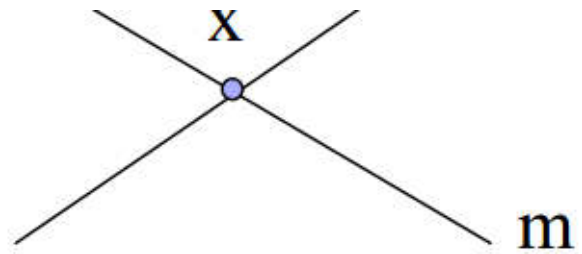
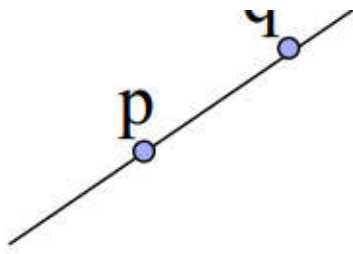
$$l = p \times q$$

也可以使用两条直线 l, m 的叉乘表示他们的交点 x

$$x = l \times m$$

见下面示例图。

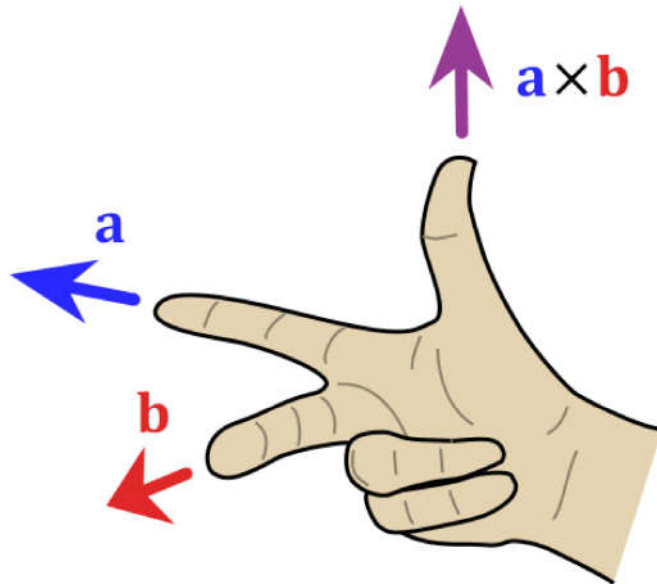




之所以可以这么简洁的表示交点是因为采用了齐次坐标的表示方式。
那么这是为什么呢？

先介绍一下叉乘（也称叉积、外积）的概念：

两个向量 a 和 b 的叉乘仅在三维空间中有定义，写作 $a \times b$
 $a \times b$ 是与向量 a, b 都垂直的向量，其方向通过右手定则（见下图）决定。

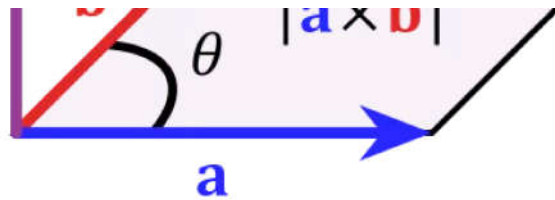


其模长等于以两个向量为边的平行四边形的面积（见下图）。
叉乘可以定义为：

$$\mathbf{a} \times \mathbf{b} = \|\mathbf{a}\| \|\mathbf{b}\| \sin(\theta) \mathbf{n}$$

其中 θ 表示 a, b 的夹角（ 0° 到 180° 之间）， $\|\mathbf{a}\|, \|\mathbf{b}\|$ 是向量 a, b 的模长
 \mathbf{n} 则是一个与向量 a, b 所构成的平面垂直的单位向量





根据叉乘定义：

向量自身叉乘结果为0，因为夹角为0。也就是说三维向量 $a \times a = 0$, $b \times b = 0$ 而点乘（也称点积，内积）的定义是

$$a \cdot b = \|a\| \cdot \|b\| \cdot \cos(\theta)$$

根据定义：如果两个向量垂直， $\cos(\theta) = 0$ ，点积也为0。

好了，经过上面点乘和叉乘定义的铺垫。下面来推导一下上面的结论：

为什么两条直线 l, m 的叉乘 $l \times m$ 等于它们的交点 p ，也就是 $p = l \times m$ ？

原因如下：首先，根据前面叉乘的定义， $l \times m$ 的结果向量（记为 $p = l \times m$ ）与 l 和 m 都垂直，根据点乘的定义，垂直的向量之间的点积为0，因此可以得到：

$$\begin{aligned} l^T * (l \times m) &= l^T * p = 0 \\ m^T * (l \times m) &= m^T * p = 0 \end{aligned}$$

因此，根据前面点在直线上的结论，可以看到 p 既在直线 l 上又在直线 m 上，所以 $p = l \times m$ 是两条直线的交点。此处 p 是齐次坐标。

同样的，可以证明，两点 p, q 的叉乘可以表示过两点的直线 l ，即 $l = p \times q$ 。（留做作业）

3 能够区分一个向量和一个点

先给出结论：

(1)从普通坐标转换成齐次坐标时

如果 (x, y, z) 是个点，则变为 $(x, y, z, 1)$;

如果 (x, y, z) 是个向量，则变为 $(x, y, z, 0)$

(2)从齐次坐标转换成普通坐标时

如果是 $(x, y, z, 1)$ ，则知道它是个点，变成 (x, y, z) ;

如果是 $(x, y, z, 0)$ ，则知道它是个向量，仍然变成 (x, y, z)

具体解释见：

<http://www.cnblogs.com/csyisong/archive/2008/12/09/1351372.html>

4 能够表达无穷远

比如 两条平行的直线 $ax+by+c=0$, $ax+by+d=0$,

可以分别用向量 $l = (a, b, c)$, $m = (a, b, d)$ 表示

根据前面直线交点的计算方法，其交点为 $l \times m$

根据叉乘计算法则

向量

$$\mathbf{u} = u_1 \mathbf{i} + u_2 \mathbf{j} + u_3 \mathbf{k}$$

$$\mathbf{v} = v_1 \mathbf{i} + v_2 \mathbf{j} + v_3 \mathbf{k}$$

的叉乘结果可以用如下方法计算得到

$$\begin{array}{l} + \mathbf{i} u_2 v_3 \\ + u_1 v_2 \mathbf{k} \\ + v_1 \mathbf{j} u_3 \\ - v_1 u_2 \mathbf{k} \\ - \mathbf{i} v_2 u_3 \\ - u_1 \mathbf{j} v_3 \end{array} \left| \begin{array}{ccc} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{array} \right|$$

最终： $l \times m = (d-c)(b, -a, 0)$ ，忽略标量 $(d-c)$ ，我们得到交点为 $(b, -a, 0)$ ，并且是齐次坐标，如果要转化为非齐次坐标，那么会得到 $(b/0, a/0)$ ，坐标是无穷大，可以认为该点为无穷远点，这与我们通常理解的：平行线相交于无穷远的概念相吻合。

因此，如果一个点的齐次坐标中，最后一个元素为0，则表示为无穷远点。

5 更简洁的表达欧氏空间变换

这是齐次坐标最重要的一个优势之一。在以后的学习中你会更加深刻的理解。

使用齐次坐标，可以方便的将加法转化为乘法，方便的表达平移。

比如我们要完成将2D坐标点 $x=[u,v]'$ 平移 $t=[t_u, t_v]$ ，如果用非齐次方法的话，是用如下的加法

$$x' \rightarrow \begin{bmatrix} u' \\ v' \end{bmatrix} = \begin{bmatrix} u + t_u \\ v + t_v \end{bmatrix} = x + t$$

如果用齐次坐标表示时可以将加法转换为乘法

$$x' \rightarrow \begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_u \\ 0 & 1 & t_v \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = Tx$$

在欧氏变换中一般有两种操作：旋转和平移。

如果我们想要将向量a进行一个标准的欧氏变换，一般是先用旋转矩阵R进行旋转，然后再用向量t进行平移，其结果 $a' = R*a + t$ ，这样看起来没什么问题。

但是，我们知道SLAM中一般都是连续的欧氏变换，所以会有多次连续的旋转和平移，假设我们将向量a进行了两次欧氏变换，分别为 R_1, t_1 和 R_2, t_2 ，分别得到：

$$b = R_1*a + t_1, c = R_2*b + t_2$$

$$\text{最终的结果 } c = R_2*(R_1*a + t_1) + t_2$$

显然，这样的变换在经过多次后会变的越来越复杂。其根本原因是上述表达方式并不是一个线性的变换关系。

此时，齐次坐标就显示出它的魅力了，如果使用齐次坐标来表达 $a' = R*a + t$ 的话可以写为：

$$\begin{bmatrix} a' \\ 1 \end{bmatrix} = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ 1 \end{bmatrix} = T \begin{bmatrix} a \\ 1 \end{bmatrix}$$

旋转和平移可以用一个矩阵T来表示，该矩阵T称为变换矩阵（transform matrix），这样欧氏变换就变成了线性关系，进行多次欧氏变换只需要连乘变换矩阵就行了，比如前面的两次欧氏变换使用齐次坐标就可以表示为：

$$\tilde{b} = T_1\tilde{a}, \tilde{c} = T_2\tilde{b} \rightarrow \tilde{c} = T_2T_1\tilde{a}$$

其中，波浪号代表齐次坐标。一般的，在SLAM中， $b = Ta$ 的形式默认都是齐次坐标。

关于齐次坐标的优势还有哪些呢？欢迎留言补充。

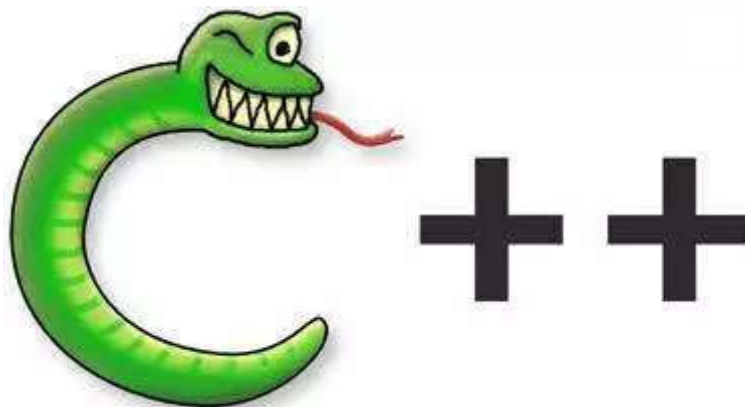
作业

证明：两点p, q 的叉乘 可以表示 过两点的直线l，即 $l = p \times q$

(提示：参考本文前面的证明)

相关阅读

从零开始一起学习SLAM | SLAM有什么用？



从零开始一起学习SLAM | C++新特性要不要学？



零基础小白，如何入门计算机视觉？