向量的线性变换与基底的线性变换

小圆滚滚

1 开宗明义

好的老师在教学当中一定会有价值观输出。很多学生可能记不住具体的知识，但是对于这些价值观的八卦，却历历在目仿若昨日。大学的时候，我线性代数没有学好，而我又觉得它好有道理。理性告诉我，要把它记住，考试要用。感性告诉我，你记不住因为你压根儿就没理解。然后我就问理性，你是什么？理性说，我就是规矩，我就是推算，只要按照我给的方法，就能得到正确的结论。然后我又问感性，你是什么？感性说，我就是观感，我就是直觉，这个人走两步我就能看出他哪里不对称、哪里不对劲儿，因为他难受的地方一定会尽量少动。

相对论特别美的地方是公平。公平的原因是线性空间的八大法则。公平的体现是相互作用力。

我特别喜欢物理，因为我总是得第一。因为我爱观察物体的运动，观察汽车、滚铁环、回旋镖我从小就能从3翼做到5翼并且精准的接回手中。我喜欢一个方块在我脑海中匀速直线运动，然后去撞击另一个方块，然后去想撞击后为什么动的是你，而不是我？

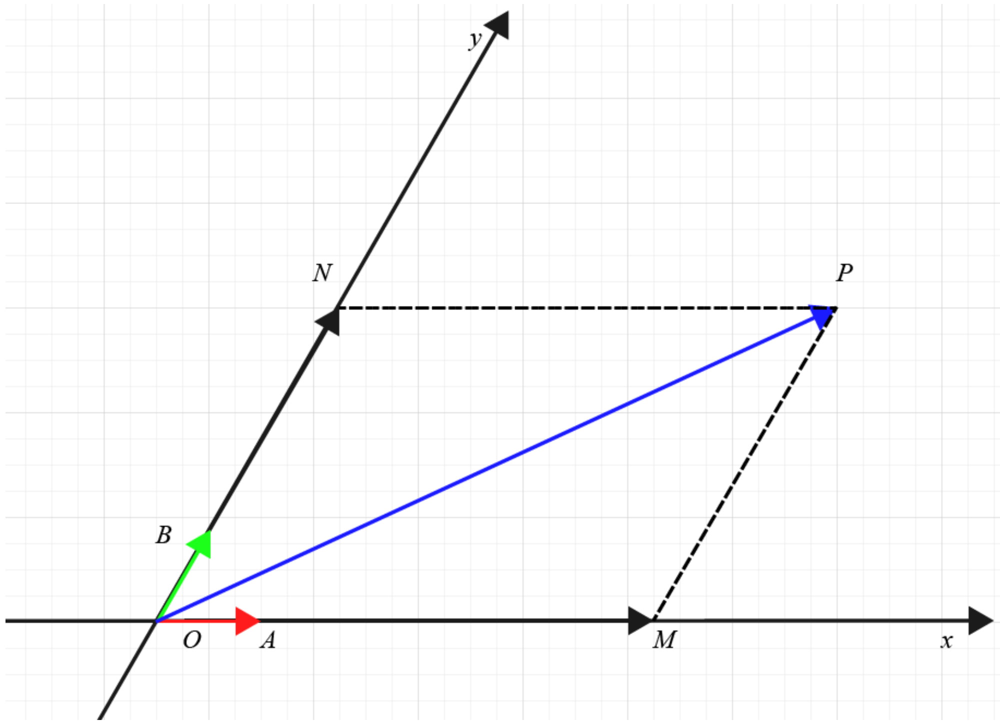
我线性代数没有学好，导致我高等数学也没有学好，这种因果关系，和我先学高等数学再学线性代数的大学排课设计没有关系、和教授的老师没有关系。因为老师很好的完成了他的任务，他们不像小学、初中的数学老师，你懂了吗？不懂？我再给你讲一遍。他们从来不问你懂不懂，或者从来不理会你的回答。然后就继续讲下去了。原因我现在也明白。他们知道有大多数人，在这么高密度的信息传输课堂上，有很多基础问题无法逐一解答。并且，除此之外，在这么大信息量或者这么大的结论输出给他们的时候，他们没打算接纳。他们也许就是来听一听进度的。他们也许就是来听一听重点方便死记硬背完成考试的。甚至他们在等待老师能不能给一个最简公式、万能公式，这辈子就在别人的模板下，舒服的活着就好。所以，老师们按照自己的节奏。雷打不动的准时准点儿、不保质但保量的完成了任务。

接下来，我会以左右逢源的形式，多角度印证、拆解。将一个概念是如何满足世界法则的，尽可能多的展现在你面前。让我们在美妙的赞叹声中，吸取大自然的精华。

首先，我们要描述一个现象。让所有人都能理解，并且没有歧义。物理学家发现，很多物理现象总结出来的规律是放之四海而皆准的。比如抛物线、悬垂线、最速降线，比如单摆、弹簧、引力。不管是在中国还是在埃及，不管是在5000年前还是在5000年后。我们的公式都可以推导出普适的物理现象。抛开时域。在不同的地域，只要参数稍作调整，就能准确计算物体的落点。最简单的就是物体的运动，运动代表变化，静止代表不变。那么不管从哪里开始运动，运动的速度、加速度规律都是适用的。那么我们选择不同的原点，就能观测物体的运动规律。描述物体运动规律、变化的最基础的量就是方向和大小。而方向是可以平移的，可以铺满整个坐标系的，但是它的运动规律不变。除了方向，就是大小。大小的计算我们经常遇到，而把方向提炼出来，又能克服平移方向带来的普适的运用原理这一层理解上的难点。那么向量这个概念的理解就算到位了。

2　向量在仿射坐标系下的运算

向量相加：



什么是矢量？

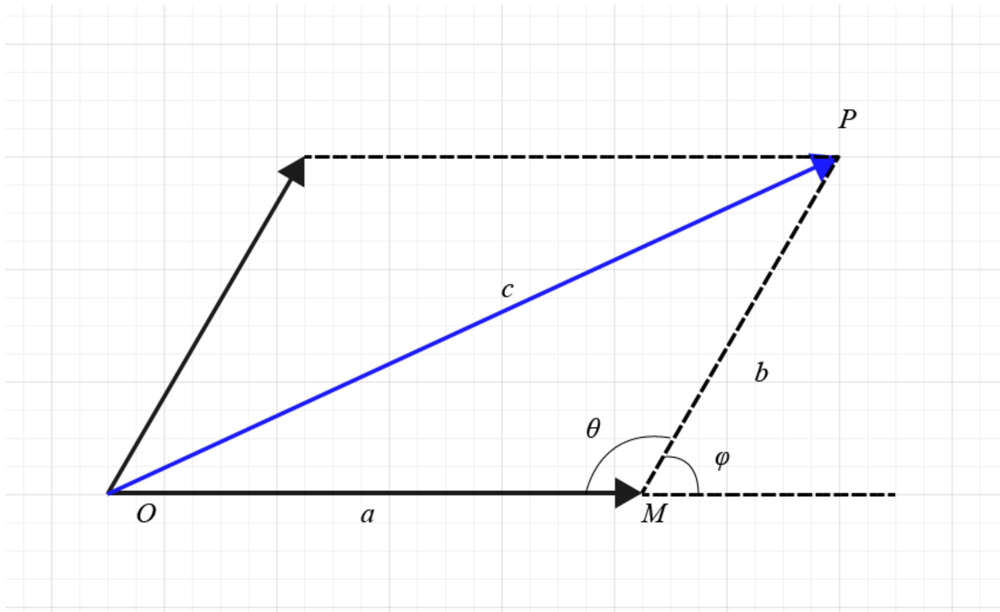
物理的本质是研究物体运动的规律，这个规律可以更换坐标系而不变，可以更换星球引力而不变。到最终，是一个世界的运行原理，放之四海而皆准。那么在空间中画一个稳固的三角形需要什么？如果按照坐标来画每一个三角形，那么就不具备通性（被钉死位置无法改变，太具体到个体了）。三角形内张面的物体的运动规律就可以由两条边确定后描述。确定的正是第三条边，我们称这第三条边就是矢量，它可以由原点共线的另外两条线段来确定。两个坐标点，四个数据，就可以确定一条有向线段，以其中一个点为开始，并且可以平移，那就是方向和大小（点向式）。这两条起始边我们称之为基底（简称基）。这两条边围成的平行四边形，可以看成是一个平面坐标系。

$\overrightarrow{OM} + \overrightarrow{MP} = \overrightarrow{OP}$

其中矢量$\overrightarrow{OA}, \overrightarrow{OB}$称为该斜坐标系的单位矢量，它们的夹角$\theta \neq 90^\circ$ ，O为坐标原点，则由平面矢量基本定理可知：对于该平面内任意给定的矢量可以表示为单位矢量的线性叠加，而且这种表达式是唯一的，即

$$\overrightarrow{OP} = x \cdot \overrightarrow{OA} + y \cdot \overrightarrow{OB} \tag{1}$$

3 余弦定理的向量证明法

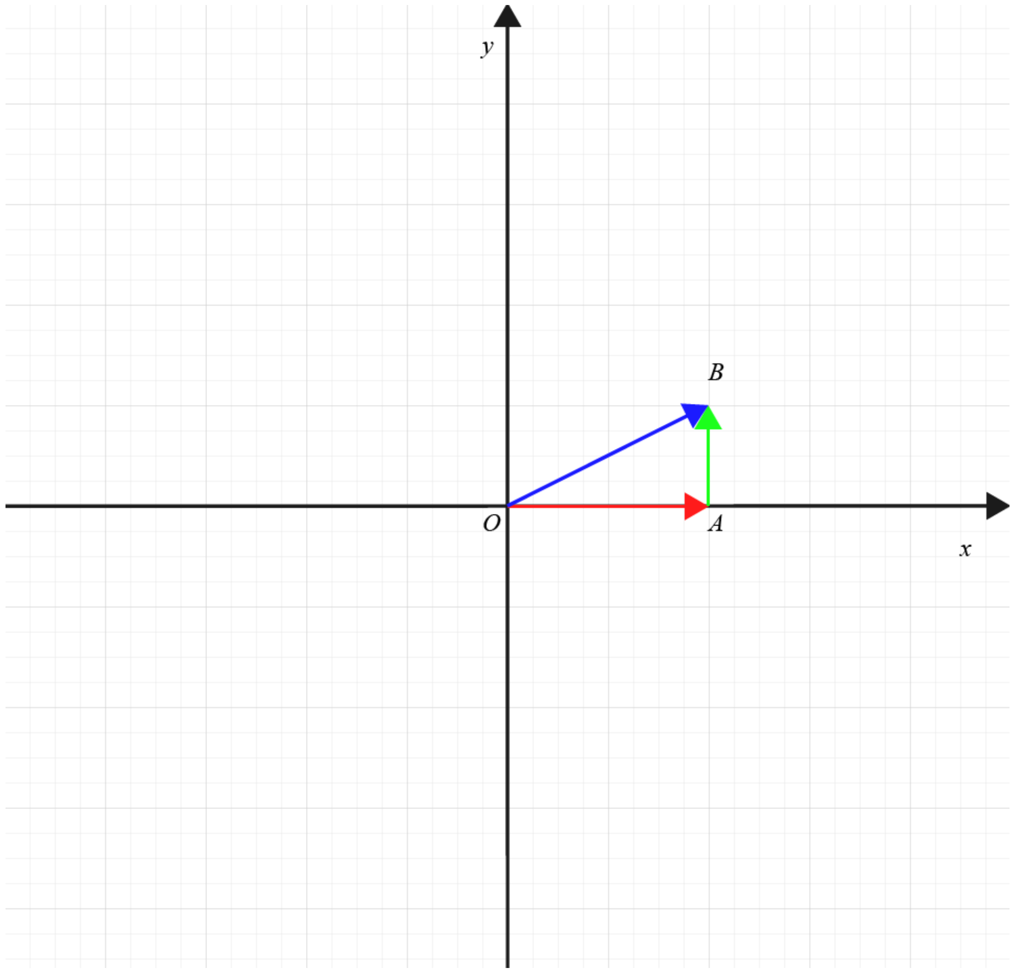


$$
\begin{aligned}
(\overrightarrow{OP})^2 &= (\overrightarrow{OM} + \overrightarrow{MP})^2 \\
&= (\overrightarrow{OM} + \overrightarrow{MP})(\overrightarrow{OM} + \overrightarrow{MP}) \\
&= (\overrightarrow{OM})^2 + (\overrightarrow{MP}^2) + 2\overrightarrow{OM}\overrightarrow{MP} \\
&= a^2 + b^2 + 2abcos < \overrightarrow{OM}, \overrightarrow{MP} > \\
&= a^2 + b^2 + 2abcos\phi \\
&= a^2 + b^2 + 2abcos(\pi - \theta) \\
&= a^2 + b^2 - 2abcos\theta
\end{aligned}
$$

其中用到向量的内积、向量的投影的双线性性概念。向量的自乘将得到自身长度的平方。是个标量。向量的相乘就是向量的内积，也是一个标量，但大小可看作一个向量在另一个向量上的投影乘以另一个向量。

注意：此处的夹角要将向量的原点对齐再看，不会超过180度。

4 向量在直角坐标系下的表示



5 矩阵的点积

矩阵的基底，相当于仿射坐标系的标架。向量（斜体加粗表示）用列矩阵（竖向）表示：

$\boldsymbol{e} = \begin{pmatrix} x \\ y \end{pmatrix}$

或者用 **转置的行矩阵** 表示： $\boldsymbol{e} = \begin{pmatrix} x & y \end{pmatrix}^T$

基底（简称基）由多个线性无关的向量组成，用矩阵的形式表示时，纵列代表一个向量，多个向量的横排构成一个方阵（中间有逗号）。 $\begin{pmatrix} \boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n \end{pmatrix}$

由定义可知，由于基底中的向量互相之间线性无关，所以两个向量$\boldsymbol{\alpha} = (x_1, y_1), \boldsymbol{\beta} = (x_2, y_2)$的点积在代数上就代表了两个向量分别分解在基底上的数值相乘，然后将乘数再相加（那意思就是在直角坐标系下）。

$\boldsymbol{\alpha} \cdot \boldsymbol{\beta} = x_1 x_2 + y_1 y_2$

更一般的：

$\boldsymbol{\alpha} \cdot \boldsymbol{\beta} = \sum\limits_{n=1}^{n} a_i b_i = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$

在几何上，两个向量$\boldsymbol{\alpha}, \boldsymbol{\beta}$，$|\boldsymbol{\alpha}|, |\boldsymbol{\beta}|$表示他们的大小，$\theta$是他们的夹角，那么内积就是

$\boldsymbol{\alpha} \cdot \boldsymbol{\beta} = |\boldsymbol{\alpha}||\boldsymbol{\beta}|cos(\boldsymbol{\alpha}, \boldsymbol{\beta})$

几何意义就是第一个向量投影到第二个向量上（这里，向量的顺序是不重要的，点积运算是可交换的），然后通过除以它们的标量长度（平方再开方）来“标准化”。这样，这个分数一定是小于等于1的，可以简单地转化成一个角度值。

当一个向量经过变换（左乘一个矩阵），生成另一个向量：

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \underbrace{\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}}_{\text{第一种解释}} = \begin{bmatrix} ax+by \\ cx+dy \end{bmatrix} = \underbrace{x \begin{bmatrix} a \\ c \end{bmatrix} + y \begin{bmatrix} b \\ d \end{bmatrix}}_{\text{第二种解释}} = \underbrace{\begin{bmatrix} \begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ \begin{bmatrix} c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \end{bmatrix}}_{\text{点积}} \tag{2}$$

第一种解释是代数的解释，就是把向量左边的部分看成一个函数（左乘一个矩阵）。

第二种解释是几何的解释，就是在新的基底下，原来的横坐标的数值在新基底的一个分向量上（横竖）分解，原来纵坐标的数值在新基底的一个分向量上（横竖）分解，然后将分解后的新数值在各个对应的部分相加。或者说，按照同等倍数去影响两个新的基底向量，这个倍数的前提就是共原点、平面线性缩放（原平面上任意一条直线在变换后不得弯曲）。参考公式(1).

第三种解释比较高深，张量的解释。相当于你用i帽、j帽$\hat{\imath}, \hat{\jmath}$（左右手手持这两根棍），张开了一个平面网，而网上粘住了一条软虫，你利用i帽、j帽这两根棍在你身体这个节点不变的原点下，进行手臂分别缩放、两臂角度张大张小、两臂顺序左右互换等操作。这条软虫会发生方向和长短的变化。这就是张量对向量产生的影响。

它满足八大法则：

1. $\boldsymbol{\alpha} + \boldsymbol{\beta} = \boldsymbol{\beta} + \boldsymbol{\alpha}$
2. $(\boldsymbol{\alpha} + \boldsymbol{\beta}) + \boldsymbol{\gamma} = \boldsymbol{\alpha} + (\boldsymbol{\beta} + \boldsymbol{\gamma})$
3. 在V中存在一个零元素$\mathbf{0}$，对任何$\boldsymbol{\alpha} \in V$，都有$\boldsymbol{\alpha} + \mathbf{0} = \boldsymbol{\alpha}$
4. 对任何$\boldsymbol{\alpha} \in V$，都有$\boldsymbol{\alpha}$的负元素$\boldsymbol{\beta} \in V$，使$\boldsymbol{\alpha} + \boldsymbol{\beta} = \mathbf{0}$
5. $1\boldsymbol{\alpha} = \boldsymbol{\alpha}$
6. $\lambda(\mu\boldsymbol{\alpha}) = (\lambda\mu)\boldsymbol{\alpha}$，常系数$\lambda\mu$
7. $(\lambda + \mu)\boldsymbol{\alpha} = \lambda\boldsymbol{\alpha} + \mu\boldsymbol{\alpha}$
8. $\lambda(\boldsymbol{\alpha} + \boldsymbol{\beta}) = \lambda\boldsymbol{\alpha} + \lambda\boldsymbol{\beta}$

前两条是加法交换律、加法结合律；规律3、5保证了加法有逆运算，利用负元素将减法转换为加法；规律6、7、8是数乘的结合律和分配率。规律5保证了非零数乘有逆运算，也就是定义了除法。

相加和数乘构成了线性变换。让两个线性无关的向量$\vec{v}, \vec{w}$带上两个标量（常系数）a, b形成新的向量$a\vec{v} + b\vec{w}$能达到平面中的每一个点。固定两个标量中的其中一个，所产生的向量的终点会描出一条直线。直线代表了方向，长度代表了大小。正好又是向量。

新基底的坐标数值是其对应的在原基底的坐标。说起来很绕，但是相当于按顺序（先i后j）、按大小、按方向（先大小再方向）重新摆放拉伸了i帽、j帽$\hat{\imath}, \hat{\jmath}$，然后可以观察到新基底下的向量对应源向量同等的发生了线性变换。

矩阵相乘，可以看成是左侧的矩阵线性变换了右侧的i帽、j帽$\hat{\imath}, \hat{\jmath}$组成的矩阵，而右侧的i帽、j帽组成的矩阵，又是要线性变换右侧还没写出来的一个向量到i帽、j帽组成的矩阵张成的空间。

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} ae+bg & af+bh \\ ce+dg & cf+dh \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e \\ g \end{bmatrix} & \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} f \\ h \end{bmatrix} \end{bmatrix}$$

拓展，矩阵相乘看成向量嵌套进矩阵相乘！若矩阵 $\boldsymbol{P}=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$

$$\boldsymbol{P}^T\cdot\boldsymbol{P}=\begin{bmatrix} a & c \\ b & d \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}=\begin{bmatrix} a\cdot a+c\cdot c & a\cdot b+c\cdot d \\ b\cdot a+d\cdot c & b\cdot b+d\cdot d \end{bmatrix}$$

令向量$\boldsymbol{\alpha}=\begin{bmatrix} a \\ c \end{bmatrix},\boldsymbol{\beta}=\begin{bmatrix} b \\ d \end{bmatrix}$代入上式

$$\boldsymbol{P}^T\cdot\boldsymbol{P}=\begin{bmatrix} \boldsymbol{\alpha}^T \\ \boldsymbol{\beta}^T \end{bmatrix}\begin{bmatrix} \boldsymbol{\alpha} & \boldsymbol{\beta} \end{bmatrix}=\begin{bmatrix} \boldsymbol{\alpha}^T\cdot\boldsymbol{\alpha} & \boldsymbol{\alpha}^T\cdot\boldsymbol{\beta} \\ \boldsymbol{\beta}^T\cdot\boldsymbol{\alpha} & \boldsymbol{\beta}^T\cdot\boldsymbol{\beta} \end{bmatrix}=\begin{bmatrix} a\cdot a+c\cdot c & a\cdot b+c\cdot d \\ b\cdot a+d\cdot c & b\cdot b+d\cdot d \end{bmatrix}$$

外部是矩阵乘法，内部是向量点积。对比公式(2)，可以领会转置的含义。

这里插一句度规的运算方法之一，就是上面这种形式，叫基向量点乘$\boldsymbol{\alpha},\boldsymbol{\beta}$。后面可以查看python源代码第四版。

而矩阵的乘法，在代数上解释为方程组如下：

$$\begin{cases} y_1=a_{11}x_1+a_{12}x_2+a_{13}x_3 \\ y_2=a_{21}x_1+a_{22}x_2+a_{23}x_3 \end{cases}$$

$$\begin{cases} x_1=b_{11}t_1+b_{12}t_2 \\ x_2=b_{21}t_1+b_{22}t_2 \\ x_3=b_{31}t_1+b_{32}t_2 \end{cases}$$

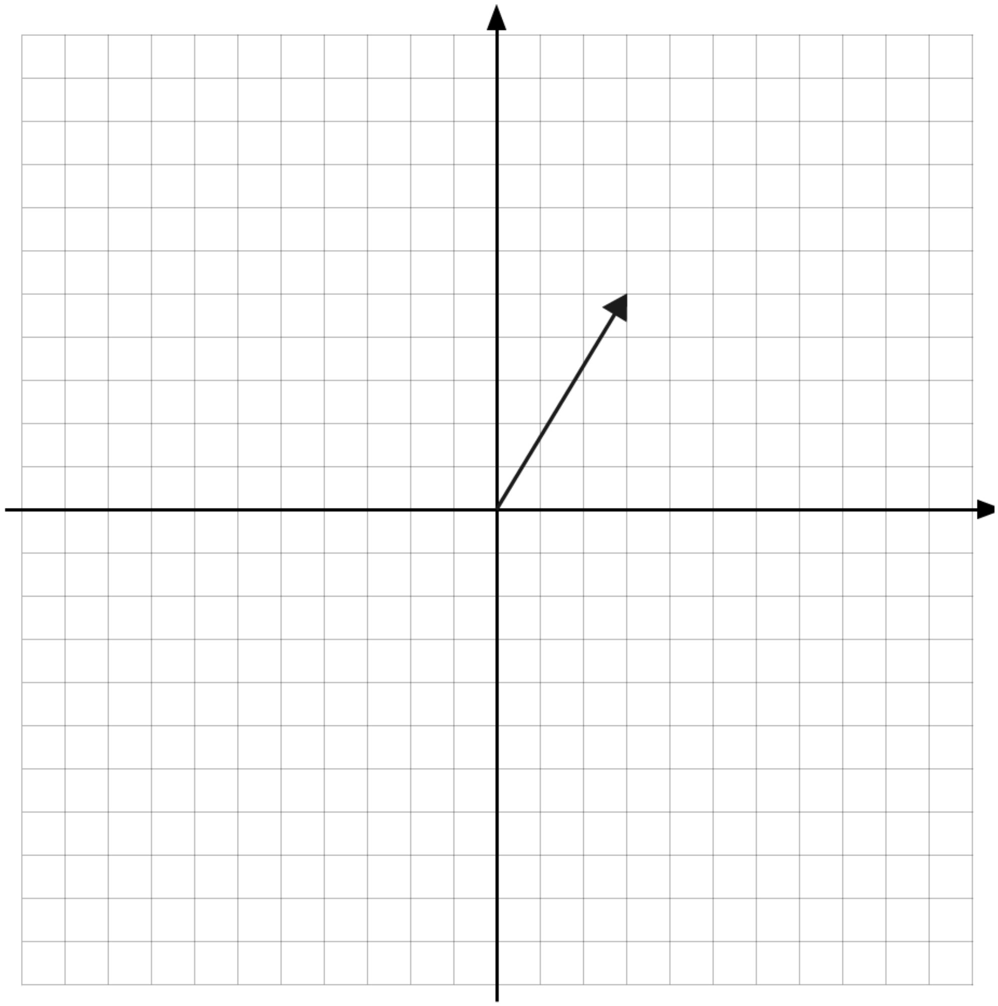
从t1，t2到y1,y2的线性变换为：$\begin{cases} y_1=(a_{11}b_{11}+a_{12}b_{21}+a_{13}b_{31})t_1+(a_{11}b_{12}+a_{12}b_{22}+a_{13}b_{32})t_2 \\ y_2=(a_{21}b_{11}+a_{22}b_{21}+a_{23}b_{31})t_1+(a_{21}b_{12}+a_{22}b_{22}+a_{23}b_{32})t_2 \end{cases}$

即 $\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{pmatrix}=\begin{pmatrix} a_{11}b_{11}+a_{12}b_{21}+a_{13}b_{31} & a_{11}b_{12}+a_{12}b_{22}+a_{13}b_{32} \\ a_{21}b_{11}+a_{22}b_{21}+a_{23}b_{31} & a_{21}b_{12}+a_{22}b_{22}+a_{23}b_{32} \end{pmatrix}$

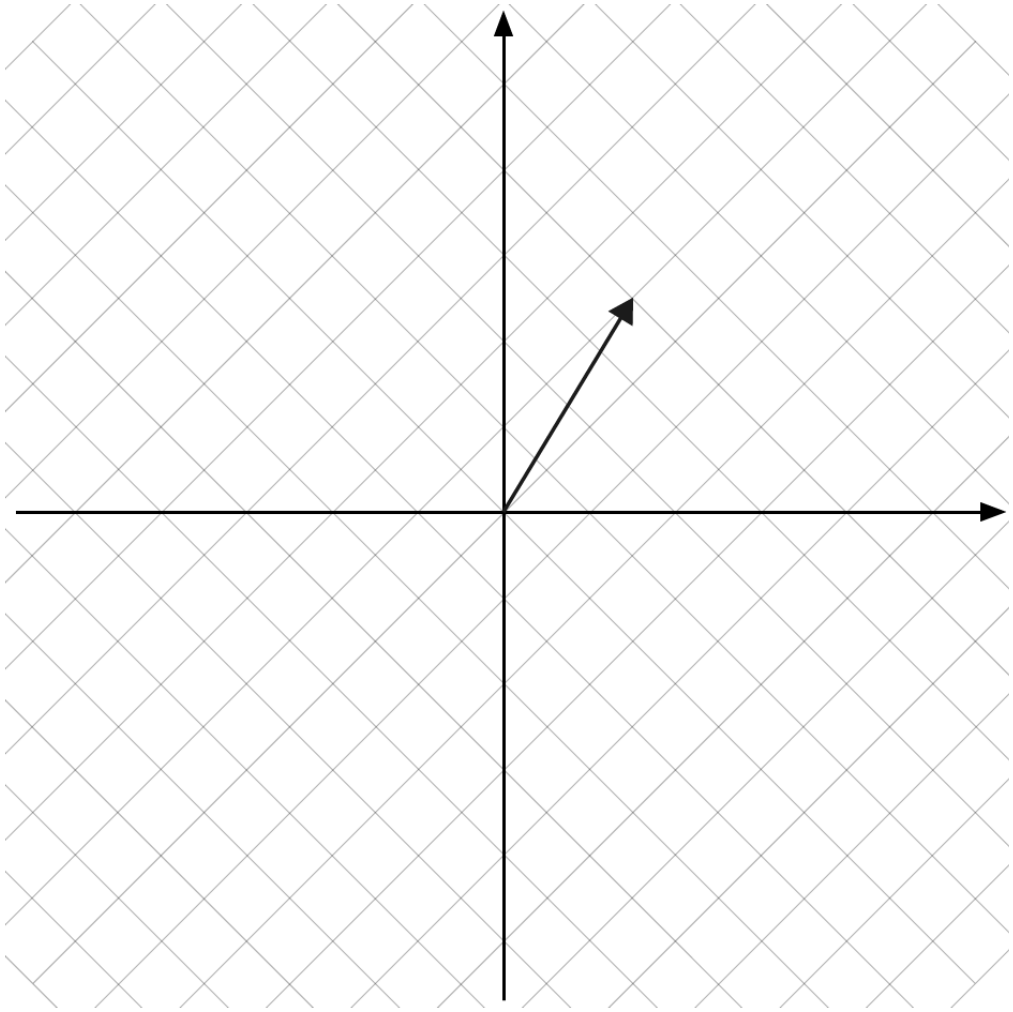
在几何思维下，可以看成是对向量（2×1的矩阵）$\begin{bmatrix} t_1 \\ t_2 \end{bmatrix}$左乘了变换矩阵$\boldsymbol{B}$（$3\times2$的矩阵），得到了三维向量$\boldsymbol{x}=\boldsymbol{Bt}$（$3\times1$的矩阵），然后又左乘了变换矩阵$\boldsymbol{y}=\boldsymbol{Ax}$，得到了二维向量$\boldsymbol{y}$。两个矩阵相乘，根据矩阵乘法的结合律，抛开针对初始向量的变化，两重线性变化可以合并成一重线性变化。于是由$\boldsymbol{t}$到$\boldsymbol{y}$的线性变化，就是两个矩阵相乘的含义。

6 不同的视角

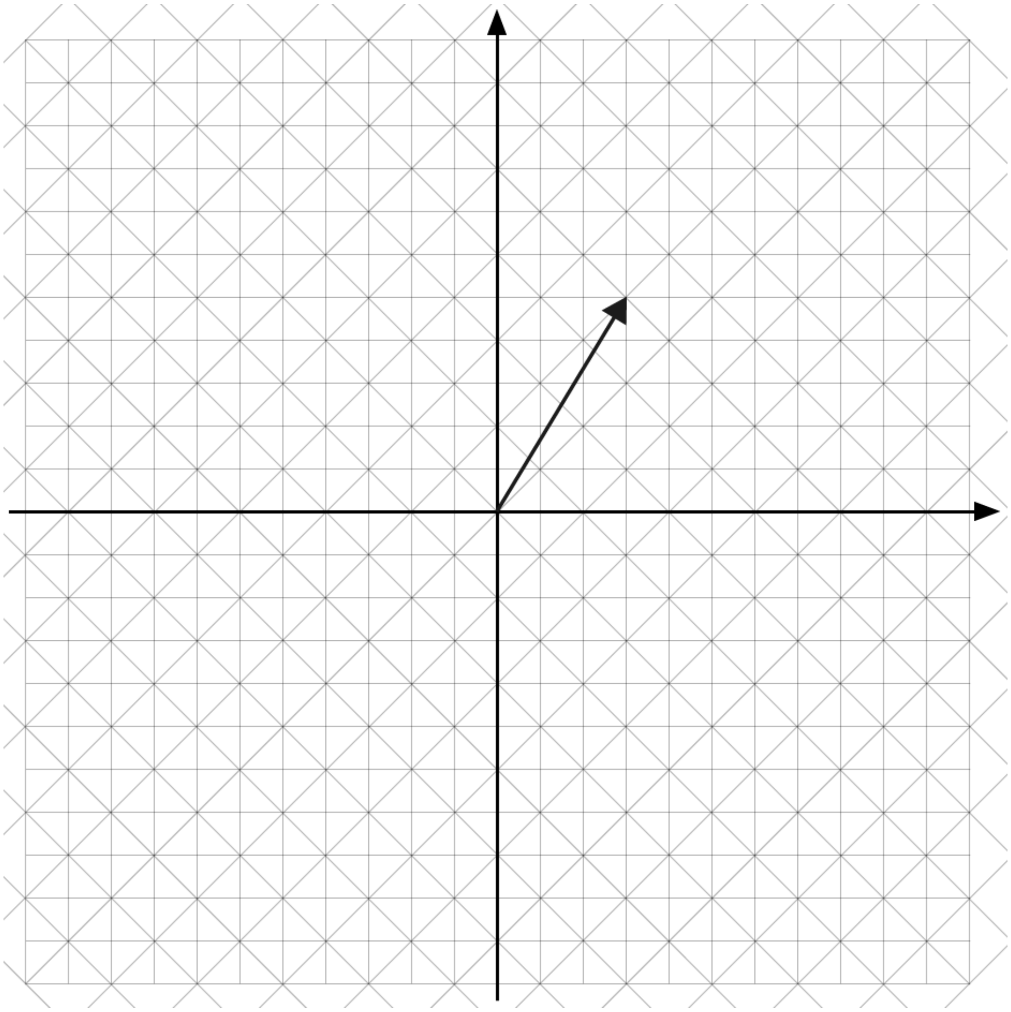
在我的视角，我首先横着走3步，然后竖着走5步。



在詹妮弗的视角，她首先横着走4步，然后竖着走1步。



我俩把各自的地图铺在一起，发现她的视角在我左前方45度，她的步子是我的1.41421倍。



但是我们都到达了同样的目的地。只不过使用了不同的坐标系、和不同的坐标。那么既然我们相等了，那么我把我们之间的联系总结一下。首先，在我看来，她的标架$(\hat{\imath},\hat{\jmath})$在我的坐标系内是这么表示的：

$$\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

用她的基变换她的向量，可以得到我的基变换我的向量，当然，我的基左乘我的向量，就是我的向量（不变）：

$$\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$$

最终都能得到在我的视角下找到的那个向量。

那么紧接着一个问题来了，如何从我的基线性变换到对方的基呢？

我们知道了对方的基在我们视角下的坐标，就可以用列向量，按$\hat{\imath},\hat{\jmath}$向量的顺序写出针对对方坐标的变换函数（左乘矩阵）。那么这个矩阵的逆矩阵，就可以将我方的坐标转换为对方的坐标。

7 线性相关

向量由方向和大小决定。方向就是平行的都是同向的。所以向量可以平移，但是向量相等。

n维空间内底线就是n个向量线性无关，多一个都是线性相关的。在此基础上，如果n-1维的向量不

能表示剩下的那一个向量，即n-1维的向量不能通过系数完成平行于剩下的那个向量这个任务，则它们线性无关。

8 左乘和右乘

向量的线性变换我们用方程组这样描述：

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n, \\ y_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n, \\ \quad \cdots \\ y_m = a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n, \end{cases}$$

相当于左乘：

$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$

其中 $\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$

基底的线性变换我们用方程组这样描述：

$$\begin{cases} \boldsymbol{\beta}_1 = p_{11}\boldsymbol{\alpha}_1 + p_{21}\boldsymbol{\alpha}_2 + \cdots + p_{n1}\boldsymbol{\alpha}_n, \\ \boldsymbol{\beta}_2 = p_{12}\boldsymbol{\alpha}_1 + p_{22}\boldsymbol{\alpha}_2 + \cdots + p_{n2}\boldsymbol{\alpha}_n, \\ \quad \cdots \\ \boldsymbol{\beta}_n = p_{1n}\boldsymbol{\alpha}_1 + p_{2n}\boldsymbol{\alpha}_2 + \cdots + p_{nn}\boldsymbol{\alpha}_n, \end{cases}$$

相当于右乘(注意这里的加法是向量的加法，并不是代数的加减)：

$\left(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n\right) = \left(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\right)\boldsymbol{P}$

其中 $\boldsymbol{P} = \begin{bmatrix} p_{11} & p_{21} & \cdots & p_{n1} \\ p_{12} & p_{22} & \cdots & p_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ p_{1n} & p_{2n} & \cdots & p_{nn} \end{bmatrix}$

注意观察P似乎是A矩阵形式的转置!!! 且是方阵。

以下证明，我们将在直角坐标系（单位正交基）下分解向量，向量相加可以看做分量相加。单看基中的一个向量

$$\boldsymbol{\beta}_1 = \begin{pmatrix} q_{11} & q_{12} & q_{13} \end{pmatrix} \begin{pmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \\ \boldsymbol{\alpha}_3 \end{pmatrix} = q_{11} \begin{bmatrix} \alpha_{1x} \\ \alpha_{1y} \\ \alpha_{1z} \end{bmatrix} + q_{12} \begin{bmatrix} \alpha_{2x} \\ \alpha_{2y} \\ \alpha_{2z} \end{bmatrix} + q_{13} \begin{bmatrix} \alpha_{3x} \\ \alpha_{3y} \\ \alpha_{3z} \end{bmatrix} = \begin{bmatrix} \beta_{1x} \\ \beta_{1y} \\ \beta_{1z} \end{bmatrix}$$

根据矩阵的转置运算规律：乘法分配率（需要交换顺序）。方程组的元素排列写成常规的列向量组，$\boldsymbol{\beta}$矩阵发生了转置，所以$\boldsymbol{Q}$矩阵转置（也就是$\boldsymbol{P} = Q^T$）就放到了原基底的右边，成了右乘。

9 基底变换

基底的变换，满足线性关系。既然都是线性变换，既然都是方阵。那么就打通了基底与坐标、基底

与基底、坐标与坐标之间的关系。

于是：对任意矢量 $\boldsymbol{A}$ ，其在旧基下存在一组坐标 $\begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$，满足

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$$

，其在新基下同样存在一组坐标 $\begin{bmatrix} A'_1 \\ A'_2 \end{bmatrix}$，满足

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{e}'_1 & \boldsymbol{e}'_2 \end{bmatrix} \begin{bmatrix} A'_1 \\ A'_2 \end{bmatrix}$$

则

$$\begin{bmatrix} \boldsymbol{e}'_1 & \boldsymbol{e}'_2 \end{bmatrix} \begin{bmatrix} A'_1 \\ A'_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$$

可得

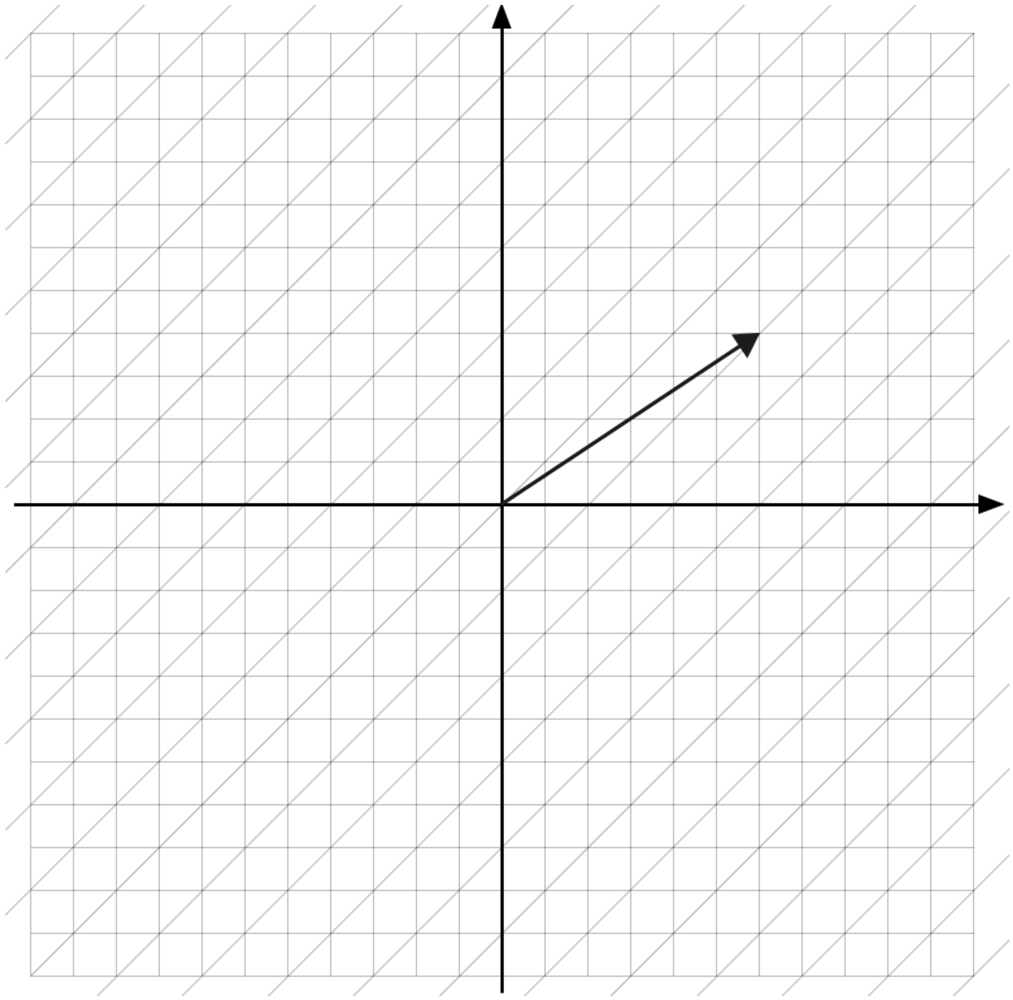
$$\begin{aligned} \begin{bmatrix} A'_1 \\ A'_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{e}'_1 & \boldsymbol{e}'_2 \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} &= (\begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 \end{bmatrix} P)^{-1} \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \\ &= P^{-1} \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \\ &= P^{-1} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \end{aligned} \tag{3}$$

以上用到了逆矩阵的乘法分配率（需要交换顺序）。

总结一下

$$\begin{cases} \begin{bmatrix} \boldsymbol{e}_1{}' & \boldsymbol{e}_2{}' \end{bmatrix} = \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 \end{bmatrix} \boldsymbol{P} \\ \begin{bmatrix} A_1{}' \\ A_2{}' \end{bmatrix} = \boldsymbol{P}^{-1} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \end{cases}$$

10 对偶基底



对于新基 $\begin{bmatrix}1 & 1\\0 & 1\end{bmatrix}$ （剪切shear，剪去左上角粘贴到右侧，面积不变）坐标 $\begin{bmatrix}2\\4\end{bmatrix}$

相当于横向走了0步，纵向走了4步。但是显然，如果按照她的地图，她竖着走的步长将是横着走的1.41421倍。增大基底的长度，读数就会变小，我们称这样的分量为逆变分量（contravariant components），用上标表示 $\begin{bmatrix}A^{1'}\\A^{2'}\end{bmatrix}$，同时，$\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$，$\{\boldsymbol{e}_1{}', \boldsymbol{e}_2{}'\}$称为协变基底（covariant basis），用下标表示。

$\boldsymbol{A} = \begin{bmatrix}\boldsymbol{e}_1 & \boldsymbol{e}_2\end{bmatrix}\begin{bmatrix}A^1\\A^2\end{bmatrix} = \begin{bmatrix}\boldsymbol{e}_1{}' & \boldsymbol{e}_2{}'\end{bmatrix}\begin{bmatrix}A^{1'}\\A^{2'}\end{bmatrix}$

有了协变基底，我们找另一个向量的垂直方向，并且和本向量夹角小于90度的向量，方向确定后，确定大小，如下：

```
>>>import math
>>>1/(math.cos(math.pi/4))
1.414213562373095
```

得到对偶基底：逆变基底（contravariant basis上标表示）。协变基底对应协变分量（covariant components下标表示）。

对空间中任意点$\boldsymbol{P}$的一组基$\{\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3\}$，我们希望在$\boldsymbol{P}$点找到另一组基$\{\boldsymbol{e}^1,\boldsymbol{e}^2,\boldsymbol{e}^3\}$，使得

$$\begin{cases}\boldsymbol{e}_i^T\boldsymbol{e}^j=0, i\neq j\\ \boldsymbol{e}_i^T\boldsymbol{e}^j=1, i=j\end{cases}$$

即

$$\begin{bmatrix}\boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3\end{bmatrix}^T\begin{bmatrix}\boldsymbol{e}^1 & \boldsymbol{e}^2 & \boldsymbol{e}^3\end{bmatrix}=\begin{bmatrix}e_{1x} & e_{1y} & e_{1z}\\ e_{2x} & e_{2y} & e_{2z}\\ e_{3x} & e_{3y} & e_{3z}\end{bmatrix}\begin{bmatrix}{e^1}_x & {e^2}_x & {e^3}_x\\ {e^1}_y & {e^2}_y & {e^3}_y\\ {e^1}_z & {e^2}_z & {e^3}_z\end{bmatrix}$$

$$=\begin{bmatrix}e_{1x}{e^1}_x+e_{1y}{e^1}_y+e_{1z}{e^1}_z & 0 & 0\\ 0 & e_{2x}{e^2}_x+e_{2y}{e^2}_y+e_{2z}{e^2}_z & 0\\ 0 & 0 & e_{3x}{e^3}_x+e_{3y}{e^3}_y+e_{3z}{e^3}_z\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$$

注意看上式的上下标对应的索引号，他们构成了矩阵的行列坐标。内涵是点积的意义：

两个前提条件，保证对偶的两个基分量是点积：

1.两组基中，有一组转置。

2.只有i=j,$cos\varphi=1\quad\varphi=0°$

点积就是用来按基的某一个基底向量来找该向量方向的分量的。逆变基底的对偶基底就是协变基底。这两个是一一对应的关系（在斜角坐标系下，尤为明显）。并且逆变基底的另一个基底是垂直于该协变基底的。这样才有克罗内克符号在夹角为90度时点积为0，才能够在点积最后的相加中去掉其他分量，唯一剩下一个分量。

那么$\{\boldsymbol{e}^1,\boldsymbol{e}^2,\boldsymbol{e}^3\}$是不是一定存在呢，答案是肯定的。

因为$\{\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3\}$线性无关，所以 $\begin{bmatrix}\boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3\end{bmatrix}$ 可逆，

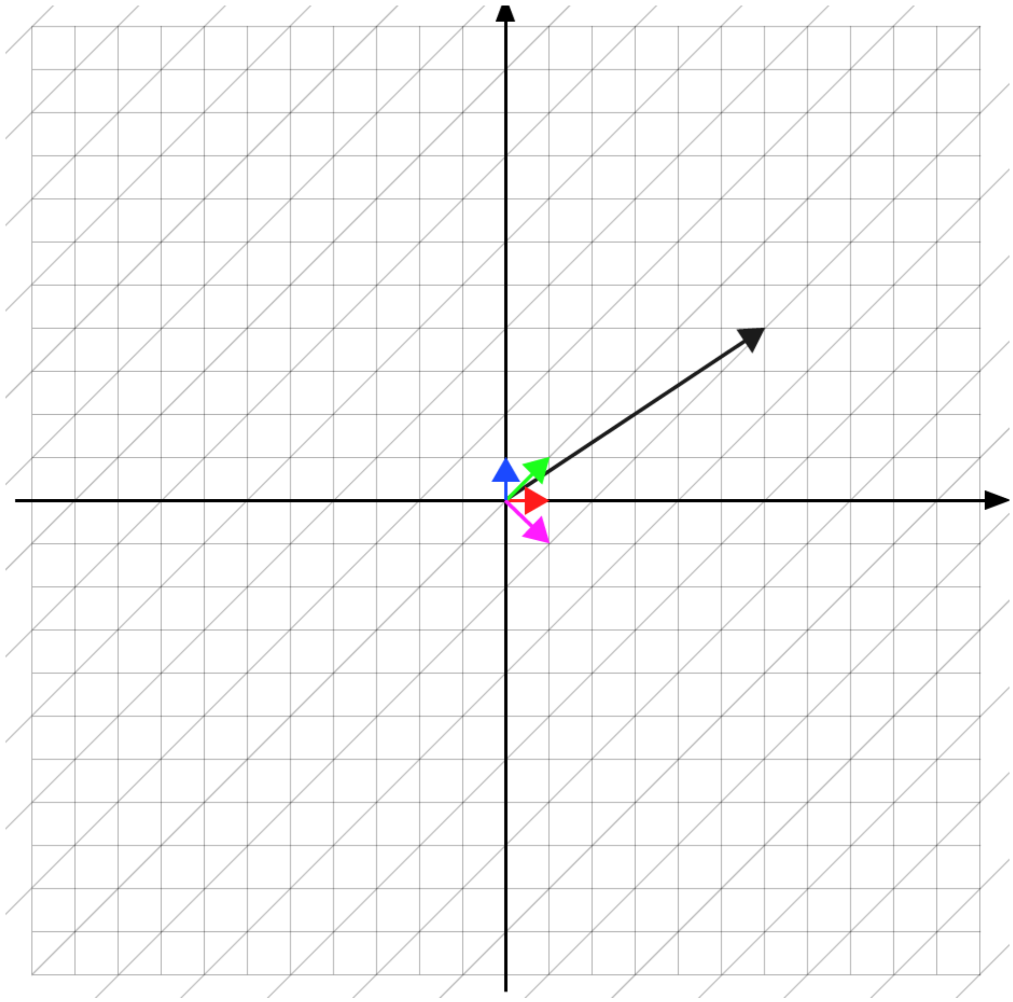
即必然存在$\{\boldsymbol{e}^1,\boldsymbol{e}^2,\boldsymbol{e}^3\}$，

且

$$\begin{bmatrix}\boldsymbol{e}^1 & \boldsymbol{e}^2 & \boldsymbol{e}^3\end{bmatrix}=(\begin{bmatrix}\boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3\end{bmatrix}^T)^{-1}\tag{4}$$

我们称$\{\boldsymbol{e}^1,\boldsymbol{e}^2,\boldsymbol{e}^3\}$就是$\{\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3\}$的**对偶基底/倒数基底**。

在斜角坐标系里，如果协变基底夹角是锐角，那么逆变基底夹角就是钝角。所以协变基底才和描述的向量以相同的方式变化（这叫协变），逆变基底和描述的向量以相反的方式变化。



图中协变基底，红色箭头是$\hat{i}$，绿色箭头是$\hat{j}$。

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 4 \end{bmatrix}=\begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

图中逆变基底，粉色箭头是$\hat{i}$，蓝色箭头是$\hat{j}$。

现在 P 点的新基$\{\boldsymbol{e}_1{}', \boldsymbol{e}_2{}', \boldsymbol{e}_3'\}$，当然也就存在新基的对偶基底$\{\boldsymbol{e}^{1\prime}, \boldsymbol{e}^{2\prime}, \boldsymbol{e}^{3\prime}\}$，根据公式(4)满足

$$\begin{bmatrix} \boldsymbol{e}^{1\prime} & \boldsymbol{e}^{2\prime} & \boldsymbol{e}^{3\prime} \end{bmatrix}=(\begin{bmatrix} \boldsymbol{e}_1{}' & \boldsymbol{e}_2{}' & \boldsymbol{e}_3{}' \end{bmatrix}^T)^{-1}=\begin{bmatrix} \boldsymbol{e}_1{}' & \boldsymbol{e}_2{}' & \boldsymbol{e}_3{}' \end{bmatrix}^{-T}$$

根据第一部分论述 $\begin{bmatrix} \boldsymbol{e}_1{}' & \boldsymbol{e}_2{}' & \boldsymbol{e}_3{}' \end{bmatrix}=\begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 \end{bmatrix}\boldsymbol{P}$ 新基可以用旧基乘以过渡矩阵表示，代入上式则：

$$\begin{bmatrix} \boldsymbol{e}^{1\prime} & \boldsymbol{e}^{2\prime} & \boldsymbol{e}^{3\prime} \end{bmatrix}=(\begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 \end{bmatrix}\boldsymbol{P})^{-T}=\begin{bmatrix} \boldsymbol{e}^1 & \boldsymbol{e}^2 & \boldsymbol{e}^3 \end{bmatrix}\boldsymbol{P}^{-T}$$

上式最后一步连续用到了转置和取逆的乘法分配率，互相抵消了。可见，从旧对偶基底到新对偶基底变换矩阵为$\boldsymbol{P}^{-T}$，其中出现了矩阵取逆操作，于是我们称对偶基底$\{\boldsymbol{e}^1, \boldsymbol{e}^2, \boldsymbol{e}^3\}$，$\{\boldsymbol{e}^{1\prime}, \boldsymbol{e}^{2\prime}, \boldsymbol{e}^{3\prime}\}$为**逆变基底（contravariant basis）**。

同时，在对偶基底$\{\boldsymbol{e}^1, \boldsymbol{e}^2, \boldsymbol{e}^3\}$，$\{\boldsymbol{e}^{1\prime}, \boldsymbol{e}^{2\prime}, \boldsymbol{e}^{3\prime}\}$下，矢量 $\boldsymbol{A}$ 的分量满足

$$\begin{bmatrix} A'_1 \\ A'_2 \\ A'_3 \end{bmatrix}=\boldsymbol{P}^T\begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix}$$

此时变换矩阵为 $\boldsymbol{P}^T$ ，无取逆操作，于是我们将矢量 $\boldsymbol{A}$ 在基底$\{\boldsymbol{e}^1, \boldsymbol{e}^2, \boldsymbol{e}^3\}$，$\{\boldsymbol{e}^{1\prime}, \boldsymbol{e}^{2\prime}, \boldsymbol{e}^{3\prime}\}$下的分

量 $\begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix}$， $\begin{bmatrix} A'_1 \\ A'_2 \\ A'_3 \end{bmatrix}$ 称为矢量 $\boldsymbol{A}$ 的**协变分量（covariant components）**。

总结一下，矢量没有所谓逆变、协变之分，其分量才有逆变、协变之分。矢量在协变基底

$\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$

下的分量为逆变分量，在逆变基底

$\{\boldsymbol{e}^1, \boldsymbol{e}^2, \boldsymbol{e}^3\}$

（协变基底的对偶基底）下的分量为协变分量。协变、逆变的概念是在变换矩阵$\boldsymbol{P}$（将旧基

$\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$

变换为新基

$\{\boldsymbol{e}_1', \boldsymbol{e}_2', \boldsymbol{e}_3'\}$

的变换矩阵）的基础上建立的。

直角坐标系下，协变基和逆变基是完全重合的，所以逆变分量（逆变量）乘以（协变基底），自然就得到了向量长度的平方。

当有两个向量相乘时，度规的稳定，可以得到系统内固定、稳定的唯一值。体现着系统的特性。

11 整个空间的坐标变换

现在上点难度。

对于一般的坐标变换

$\begin{cases} x_1 = x_1(u_1, u_2, u_3) \\ x_2 = x_2(u_1, u_2, u_3) \\ x_3 = x_3(u_1, u_2, u_3) \end{cases}$

我们简记为

$\mathbf{x} = \mathbf{x}(\mathbf{u})$

则对于任意点P的旧基$\mathbf{X}$、新基$\mathbf{U}$ ，有

$\mathbf{U} = \frac{\partial \mathbf{R}}{\partial \mathbf{u}} = \frac{\partial \mathbf{R}}{\partial \mathbf{x}} \frac{\partial \mathbf{x}}{\partial \mathbf{u}} = \mathbf{X} \frac{\partial \mathbf{x}}{\partial \mathbf{u}}$

记变换矩阵 $\frac{\partial \mathbf{x}}{\partial \mathbf{u}} = \mathbf{T}$为讨论“协变”、“逆变”概念的基准。

对于微分位移矢量$d\mathbf{x}$ 和 $d\mathbf{u}$，有

$d\mathbf{u} = \frac{\partial \mathbf{u}}{\partial \mathbf{x}} d\mathbf{x} = (\frac{\partial \mathbf{x}}{\partial \mathbf{u}})^{-1} d\mathbf{x} = \mathbf{T}^{-1} d\mathbf{x}$

显然，微分位移矢量的变换遵循“逆变”规则。

对于梯度矢量 $\frac{\partial f}{\partial \mathbf{x}}$ 和 $\frac{\partial f}{\partial \mathbf{u}}$ ，有

$\frac{\partial f}{\partial \mathbf{u}} = \frac{\partial f}{\partial \mathbf{x}} \frac{\partial \mathbf{x}}{\partial \mathbf{u}} = \frac{\partial f}{\partial \mathbf{x}} \mathbf{T}$

显然，梯度矢量的变换遵循“协变”规则。

对于微分位移矢量的模长的平方 $d^2 s$,有

$d^2 s = (d\mathbf{x})^T \mathbf{G_X} d\mathbf{x} = (\frac{\partial \mathbf{x}}{\partial \mathbf{u}} d\mathbf{u})^T \mathbf{G_X} (\frac{\partial \mathbf{x}}{\partial \mathbf{u}} d\mathbf{u})$
$= (d\mathbf{u})^T (\frac{\partial \mathbf{x}}{\partial \mathbf{u}})^T \mathbf{G_X} \frac{\partial \mathbf{x}}{\partial \mathbf{u}} d\mathbf{u} = (d\mathbf{u})^T (\mathbf{T})^T \mathbf{G_X} \mathbf{T} d\mathbf{u}$
$= (d\mathbf{u})^T \mathbf{G_U} d\mathbf{u}$

即旧基$\mathbf{X}$下的度规 $\mathbf{G_X}$ 、新基$\mathbf{U}$下的度规 $\mathbf{G_U}$ 满足关系式

$\mathbf{G_U} = (\frac{\partial \mathbf{x}}{\partial \mathbf{u}})^T \mathbf{G_X} \frac{\partial \mathbf{x}}{\partial \mathbf{u}} = (\mathbf{T})^T \mathbf{G_X} \mathbf{T}$

显然，度规（二阶张量）的变换遵循“协变”规则。

运行程序可以看出协变和逆变在基底和分量上的倍数关系。

12 长度的计算

```
请在两行内输入一个2阶矩阵(列元素代表了新基在直角坐标系下的坐标):
1 1
0 1
协变基底是:
1.0 1.0
0.0 1.0
逆变基底是:
1.0 -0.0
-1.0 1.0
请分两行输入1个坐标:
6
4
逆变分量为（2.000 4.000）
协变分量为（6.000 10.000）
逆变分量转换为协变分量的协变度规张量:
1.0 1.0
1.0 2.0
协变分量转换为逆变分量的逆变度规张量:
2.0 -1.0
-1.0 1.0
```

由于我们选取的起始基底是单位正交基，单位矩阵左乘或者右乘其他矩阵，都会保持不变，所以过渡矩阵P，实际上就可以看成是新基底的$\hat{\imath},\hat{\imath}$帽向量横向排成的矩阵。

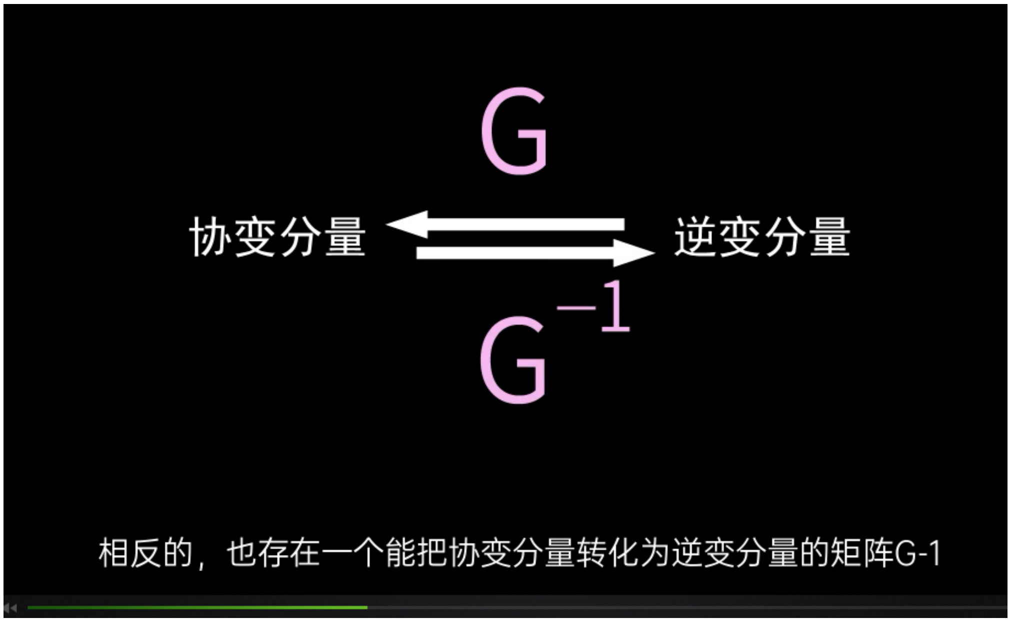
$$\because \begin{bmatrix} A^{1\prime} \\ A^{2\prime} \end{bmatrix} = \boldsymbol{P}^{-1} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$$

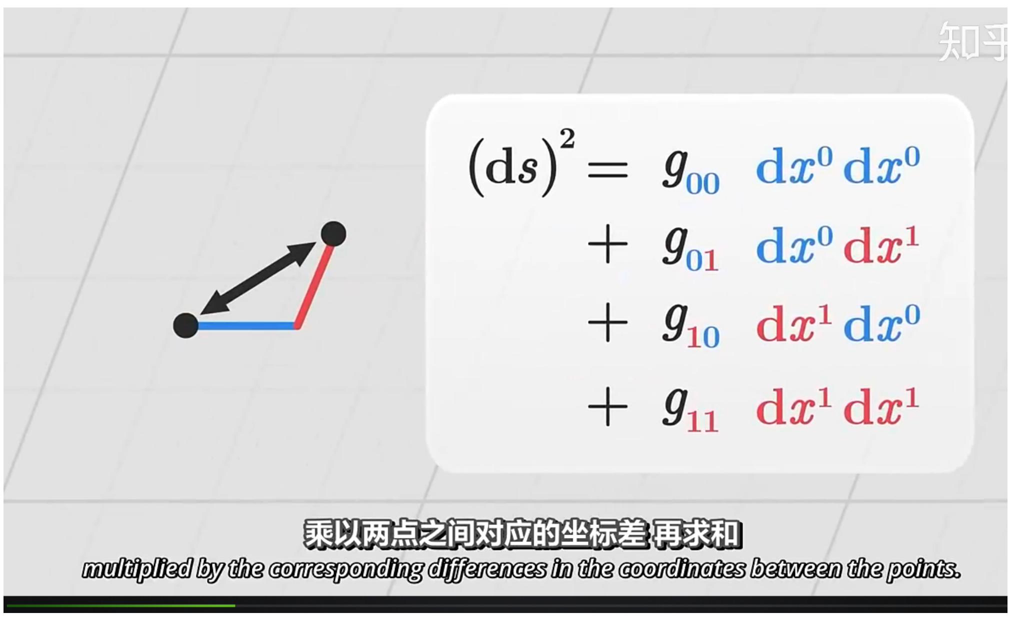
$$\begin{bmatrix} A_1{}' \\ A_2{}' \end{bmatrix} = \boldsymbol{P}^{T} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$$

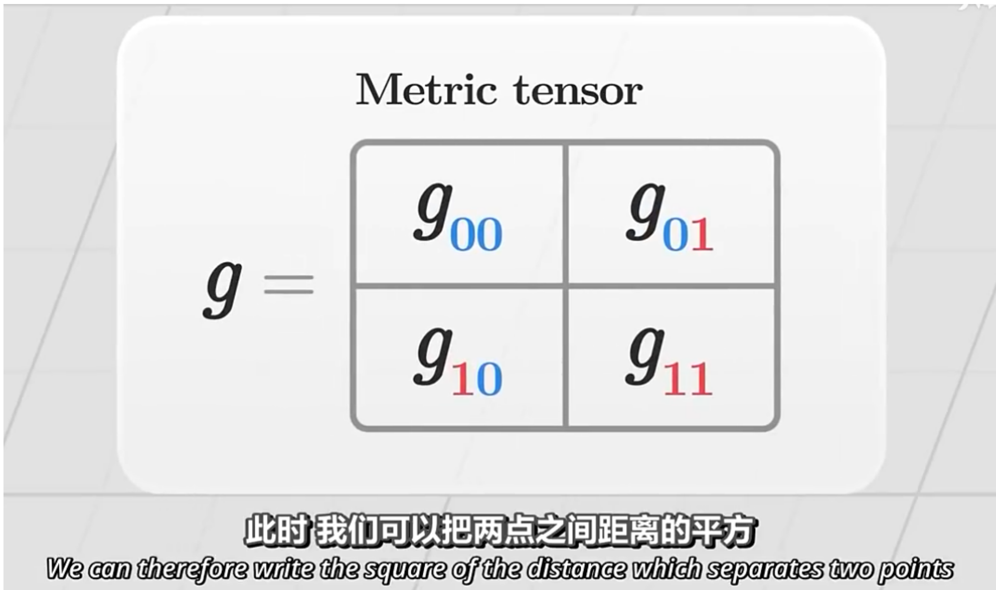
$$\therefore \boldsymbol{G} = \boldsymbol{P}^T \cdot \boldsymbol{P}$$

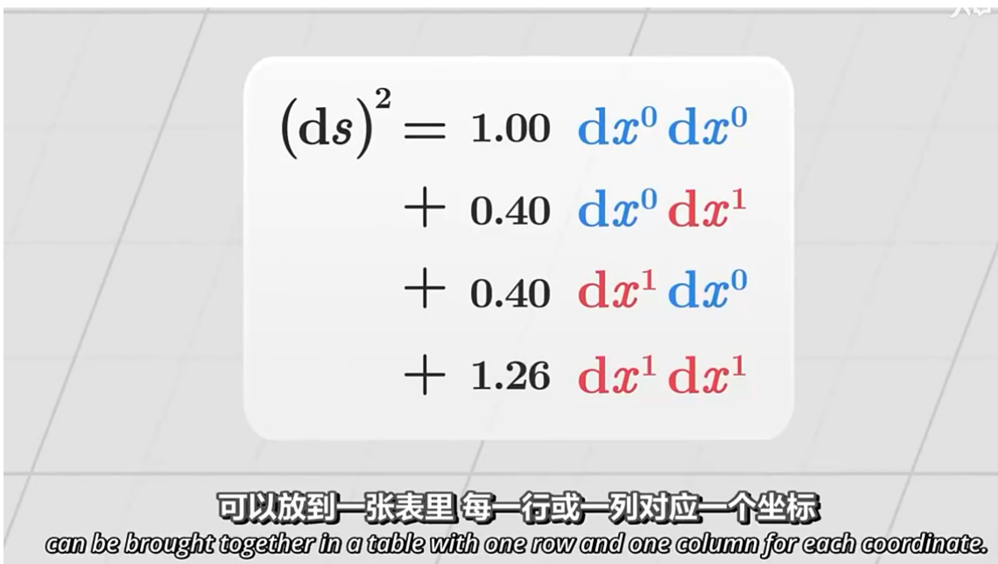
简单总结就是，协变度规张量就是基变换矩阵（过渡矩阵）的转置与自己的乘积形成的矩阵。

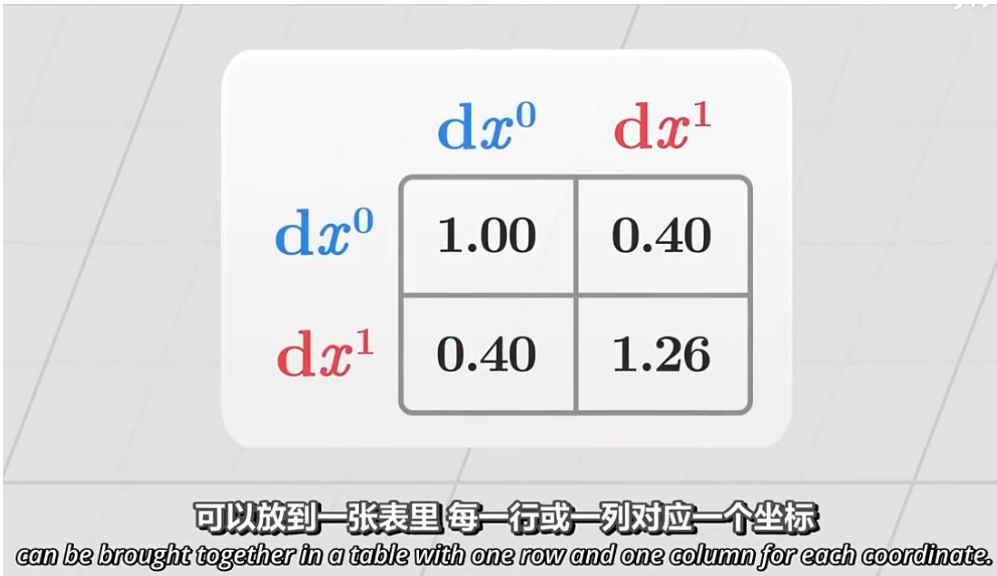
$$\begin{aligned}
(ds)^2 &= g_{\mu\nu} dx^\mu dx^\nu \\
&= 1 * 2 * 2 \\
&+ 1 * 2 * 4 \\
&+ 1 * 4 * 2 \\
&+ 2 * 4 * 4 \\
&= 52 = 6^2 + 4^2
\end{aligned}$$


G
协变分量
逆变分量
G^{-1}
相反的，也存在一个能把协变分量转化为逆变分量的矩阵G-1


知乎
$(\mathrm{d}s)^2 = g_{00}\,\mathrm{d}x^0\,\mathrm{d}x^0 + g_{01}\,\mathrm{d}x^0\,\mathrm{d}x^1 + g_{10}\,\mathrm{d}x^1\,\mathrm{d}x^0 + g_{11}\,\mathrm{d}x^1\,\mathrm{d}x^1$
乘以两点之间对应的坐标差再求和
multiplied by the corresponding differences in the coordinates between the points.


Metric tensor
$g = \begin{array}{|c|c|} \hline g_{00} & g_{01} \\ \hline g_{10} & g_{11} \\ \hline \end{array}$
此时我们可以把两点之间距离的平方
We can therefore write the square of the distance which separates two points


$(\mathrm{d}s)^2 = 1.00\ \mathrm{d}x^0\,\mathrm{d}x^0$
$+\ 0.40\ \mathrm{d}x^0\,\mathrm{d}x^1$
$+\ 0.40\ \mathrm{d}x^1\,\mathrm{d}x^0$
$+\ 1.26\ \mathrm{d}x^1\,\mathrm{d}x^1$
可以放到一张表里 每一行或一列对应一个坐标
can be brought together in a table with one row and one column for each coordinate.


dx^0 dx^1
dx^0 1.00 0.40
dx^1 0.40 1.26
可以放到一张表里 每一行或一列对应一个坐标
can be brought together in a table with one row and one column for each coordinate.