

1.2 长度与点积

1 $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ 与 $\mathbf{w} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$ 的“点积”为 $\mathbf{v} \cdot \mathbf{w} = (1)(4) + (2)(5) = 4 + 10 = 14$ 。

2 因为 $\mathbf{v} \cdot \mathbf{w}$ 是 0: $(1)(4) + (3)(-4) + (2)(4) = 0$, 所以 $\mathbf{v} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$ 与 $\mathbf{w} = \begin{bmatrix} 4 \\ -4 \\ 4 \end{bmatrix}$ 是垂直的。

3 $\mathbf{v} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$ 长度的平方为 $\mathbf{v} \cdot \mathbf{v} = 1 + 9 + 4 = 14$ 。其长度为 $\|\mathbf{v}\| = \sqrt{14}$ 。

4 于是 $\mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|} = \frac{\mathbf{v}}{\sqrt{14}} = \frac{1}{\sqrt{14}} \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$ 的长度为 $\|\mathbf{u}\| = 1$ 。检验 $\frac{1}{14} + \frac{9}{14} + \frac{4}{14} = 1$ 。

5 \mathbf{v} 与 \mathbf{w} 的夹角 θ 有 $\cos \theta = \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\| \|\mathbf{w}\|}$ 。

6 $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ 与 $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ 的夹角有 $\cos \theta = \frac{1}{(1)(\sqrt{2})}$ 。该角为 $\theta = 45^\circ$ 。

7 所有角都有 $|\cos \theta| \leq 1$ 。因此, 所有向量都有 $|\mathbf{v} \cdot \mathbf{w}| \leq \|\mathbf{v}\| \|\mathbf{w}\|$ 。

第一节放弃了讲向量相乘。现在我们继续来定义 \mathbf{v} 与 \mathbf{w} 的“点积”。这个乘法包含单独的积 $v_1 w_1$ 和 $v_2 w_2$, 但它并不止于此。这两个数加起来得出一个数 $\mathbf{v} \cdot \mathbf{w}$ 。

以下是几何部分 (向量长度及它们夹角的余弦)。

$\mathbf{v} = (v_1, v_2)$ 与 $\mathbf{w} = (w_1, w_2)$ 的点积或者说内积是数 $\mathbf{v} \cdot \mathbf{w}$:

$$\mathbf{v} \cdot \mathbf{w} = v_1 w_1 + v_2 w_2. \quad (1)$$

例 1 向量 $\mathbf{v} = (4, 2)$ 与 $\mathbf{w} = (-1, 2)$ 点积为零:

点积为 0
垂直向量

$$\begin{bmatrix} 4 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 2 \end{bmatrix} = -4 + 4 = 0.$$

在数学中, 0 总是一个特别的数。对于点积, 它意味着这两个向量是垂直的。它们的夹角是 90° 。当我们在图 1.1 中画出它们时, 我们见到了一个矩形 (不仅仅是任一平行四边形)。垂直向量最清晰的例子是沿 x 轴的 $\mathbf{i} = (1, 0)$ 与沿 y 轴向上的 $\mathbf{j} = (0, 1)$ 。再一次地, 点积为 $\mathbf{i} \cdot \mathbf{j} = 0 + 0 = 0$ 。向量 \mathbf{i} 与 \mathbf{j} 呈直角。

$v = (1, 2)$ 与 $w = (3, 1)$ 的点积是 5。马上, $v \cdot w$ 将揭露 v 与 w 的夹角 (非 90°)。请检验 $w \cdot v$ 也是 5。

点积 $w \cdot v$ 等于 $v \cdot w$ 。 v 与 w 的顺序没有差别。

例 2 在点 $x = -1$ (0 的左边) 置权重 4 且在点 $x = 2$ (0 的右边) 置权重 2。 x 轴将在中点平衡 (就像拉锯)。因为其点积为 $(4)(-1) + (2)(2) = 0$, 所以权重平衡。

这是一个典型工程与科学的例子。权重向量是 $(w_1, w_2) = (4, 2)$ 。距中点的距离向量为 $(v_1, v_2) = (-1, 2)$ 。权重乘以距离, w_1v_1 及 w_2v_2 , 得出“力矩”。跷跷板平衡的方程为 $w_1v_1 + w_2v_2 = 0$ 。

例 3 点积投入到经济与商业中。我们有三种商品要购买与销售。它们的单价为 (p_1, p_2, p_3) ——这是“价格向量” p 。我们购买或销售的数量为 (q_1, q_2, q_3) ——当我们售出时为正、购入时为负。以价格 p_1 售出 q_1 个赚取 q_1p_1 。总收益 (数量 q 乘以价格 p) 是**三维点积 $p \cdot q$** ：

$$\text{收益} = (q_1, q_2, q_3) \cdot (p_1, p_2, p_3) = q_1p_1 + q_2p_2 + q_3p_3 = \text{点积}。$$

零点积意味着“账目平衡”。若 $q \cdot p = 0$ 则售出总计等于购入总计。于是 p 与 q 垂直 (在三维空间)。一家有上千种商品的超市很快就进入高维度。

小注记：电子表格在管理中变得至关重要。它们计算线性组合与点积。你在屏幕上见到的是个矩阵。

重点 对于 $v \cdot w$, 每个 v_i 乘以 w_i 。因此 $v \cdot w = v_1w_1 + \cdots + v_nw_n$ 。

长度与单位向量

一个重要实例是向量与它自己的点积。在此情况下 v 等于 w 。当向量是 $v = (1, 2, 3)$ 时, 与它自己的点积为 $v \cdot v = \|v\|^2 = 14$ ：

$$\begin{array}{l} \text{点积 } v \cdot v \\ \text{长度平方} \end{array} \quad \|v\|^2 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = 1 + 4 + 9 = 14$$

向量夹角不是 90° 而是 0° 。因为 v 不与它自己垂直, 所以答案不是 0。点积 $v \cdot v$ 得出 v 长度的平方。

定义 向量 v 的长度 $\|v\|$ 是 $v \cdot v$ 的平方根：

$$\text{长度} = \|v\| = \sqrt{v \cdot v} = (v_1^2 + v_2^2 + \cdots + v_n^2)^{1/2}。$$

在二维中长度是 $\sqrt{v_1^2 + v_2^2}$ 。在三维中它是 $\sqrt{v_1^2 + v_2^2 + v_3^2}$ 。通过上面的计算, $v = (1, 2, 3)$ 的长度为 $\|v\| = \sqrt{14}$ 。

这里 $\|v\| = \sqrt{v \cdot v}$ 就是表示向量的箭头的普通长度。假设分量为 1 和 2, 则箭头是直角三角形的第 3 条边 (图 1.6)。毕达哥拉斯定理 $a^2 + b^2 = c^2$ 关联着三条边: $1^2 + 2^2 = \|v\|^2$ 。

对于 $v = (1, 2, 3)$ 的长度, 我们两次运用直角定理。基向量 $(1, 2, 0)$ 的长度为 $\sqrt{5}$ 。这个基向量与直立向上的 $(0, 0, 3)$ 垂直。因此这个盒子的对角线长度为 $\|v\| = \sqrt{5 + 9} = \sqrt{14}$ 。

四维向量的长度应当是 $\sqrt{v_1^2 + v_2^2 + v_3^2 + v_4^2}$ 。于是向量 $(1, 1, 1, 1)$ 长度为 $\sqrt{1^2 + 1^2 + 1^2 + 1^2} = 2$ 。这是四维空间中穿过单位立方体的对角线。 n 维立方体的对角线长度为 \sqrt{n} 。

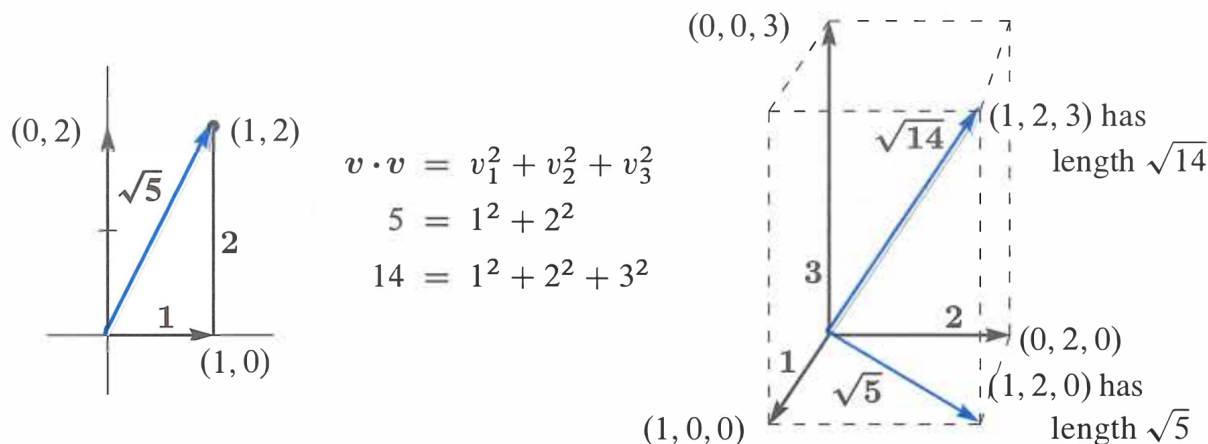


图 1.6: 二维与三维向量长度 $\sqrt{\mathbf{v} \cdot \mathbf{v}}$ 。

“单位”一词总是说明一些度量值等于“1”。单位价格是一件的价格。一个立方体的边长为 1。一个单位圆是半径为 1 的圆。现在我们理解了“单位向量”的含义。

定义 一个单位向量 \mathbf{u} 是一个长度等于 1 的向量。因此 $\mathbf{u} \cdot \mathbf{u} = 1$ 。

一个四维例子是 $\mathbf{u} = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ 。于是 $\mathbf{u} \cdot \mathbf{u}$ 是 $\frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = 1$ 。我们将 $\mathbf{v} = (1, 1, 1, 1)$ 除以它的长度 $\|\mathbf{v}\| = 2$ 以得出这个单位向量。

例 4 沿 x 轴和 y 轴的单位向量写为 \mathbf{i} 和 \mathbf{j} 。在 xy 平面，与 x 轴成“ θ ”角的单位向量是 $(\cos \theta, \sin \theta)$ ：

$$\text{单位向量 } \mathbf{i} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{与} \quad \mathbf{j} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{与} \quad \mathbf{u} = \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix}.$$

当 $\theta = 0$ 时，水平向量 \mathbf{u} 为 \mathbf{i} 。当 $\theta = 90^\circ$ 时，竖直向量为 \mathbf{j} 。因为 $\cos^2 \theta + \sin^2 \theta = 1$ ，所以在任何角度，其分量 $\cos \theta$ 和 $\sin \theta$ 都得出 $\mathbf{u} \cdot \mathbf{u} = 1$ 。这些向量延伸为图 1.7 中的单位圆。因此 $\cos \theta$ 和 $\sin \theta$ 仅仅是单位圆在 θ 角上的坐标。

由于 $(2, 2, 1)$ 长度为 3，向量 $(\frac{2}{3}, \frac{2}{3}, \frac{1}{3})$ 长度为 1。检验 $\mathbf{u} \cdot \mathbf{u} = \frac{4}{9} + \frac{4}{9} + \frac{1}{9} = 1$ 。为了得单位向量，将任意非零向量 \mathbf{v} 除以它的长度 $\|\mathbf{v}\|$ 。

单位向量 $\mathbf{u} = \mathbf{v} / \|\mathbf{v}\|$ 是与 \mathbf{v} 同方向的单位向量。

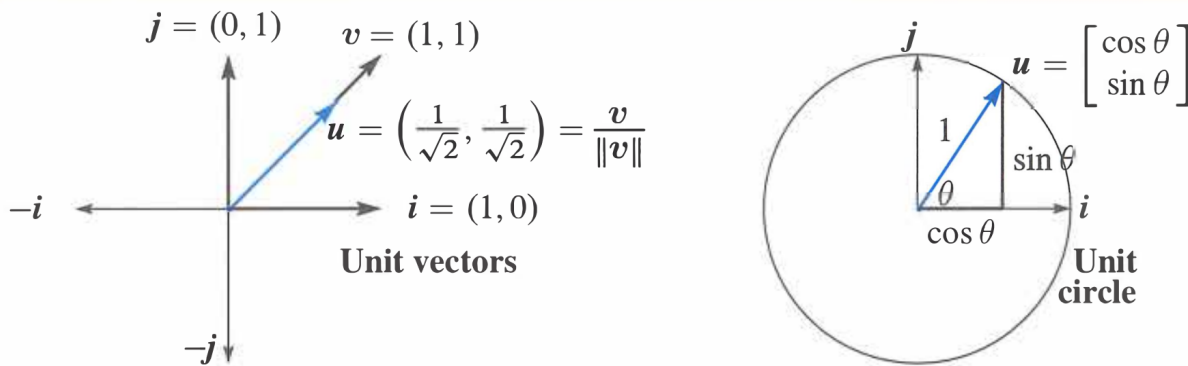


图 1.7: 坐标向量 i 与 j 。 $v = (1, 1)$ 除以它的长度 $\|v\| = \sqrt{2}$ 为 (左边) 45° 角处的单位向量 u 。单位向量 $u = (\cos \theta, \sin \theta)$ 在角 θ 处。

两向量的夹角

我们阐明了垂直向量具有 $v \cdot w = 0$ 。当夹角为 90° 时，其点积为 0。为了解释这一点，我们得将角度与点积联系起来。然后我们演示 $v \cdot w$ 如何求出任意两非零向量的夹角。

直角 当 v 与 w 垂直时，其点积为 $v \cdot w = 0$ 。

证明 当 v 与 w 垂直时，它们构成直角三角形的两条边。第三条边为 $v - w$ (图 1.8 中横跨着的斜边)。关于直角三角形边的毕达哥拉斯定理是 $a^2 + b^2 = c^2$:

$$\text{垂直向量} \quad \|v\|^2 + \|w\|^2 = \|v - w\|^2 \quad (2)$$

在二维中写出这些长度的公式，此方程为

$$\text{毕达哥拉斯} \quad (v_1^2 + v_2^2) + (w_1^2 + w_2^2) = (v_1 - w_1)^2 + (v_2 - w_2)^2. \quad (3)$$

右边以 $v_1^2 - 2v_1w_1 + w_1^2$ 开始。于是方程两边都有 v_1^2 与 w_1^2 然后它们消去了，剩下 $-2v_1w_1$ 。 v_2^2 与 w_2^2 也消去了，剩下 $-2v_2w_2$ 。(在三维中应当有 $-2v_3w_3$ 。) 立即除以 -2 来看出 $v \cdot w = 0$ (原书此处有误!):

$$0 = -2v_1w_1 - 2v_2w_2 \quad \text{推出} \quad v_1w_1 + v_2w_2 = 0 \quad (4)$$

结论 直角产生 $v \cdot w = 0$ 。当夹角为 $\theta = 90^\circ$ 时，点积为 0。那时 $\cos \theta = 0$ 。因为 $0 \cdot w$ 总是 0，所以零向量 $v = 0$ 与每个向量 w 都垂直。

现假设 $v \cdot w$ 非零。它可能是正的也可能是负的。 $v \cdot w$ 的符号会立即向我们表明是小于直角还是大于直角。当 $v \cdot w$ 为正时，夹角小于 90° 。当 $v \cdot w$ 为负时，夹角大于 90° 。图 1.8 的右边展示了一个典型的向量 $v = (3, 1)$ 。因为 $v \cdot w = 6$ 为正，所以其与 $w = (1, 3)$ 的夹角小于 90° 。

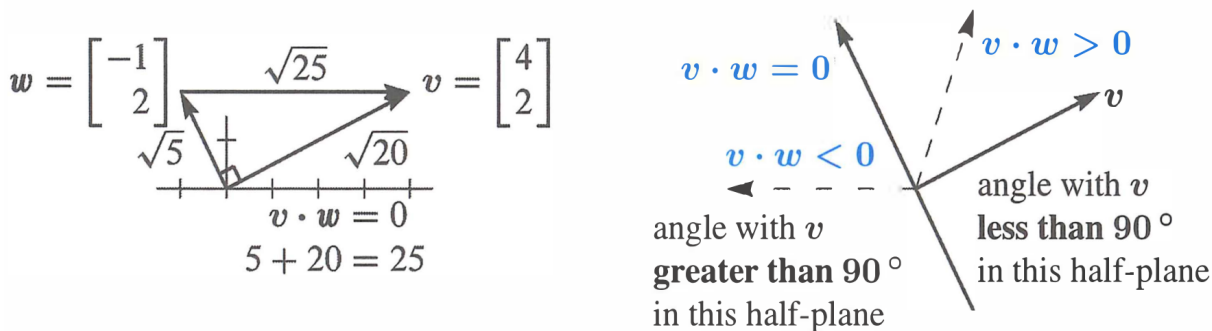


图 1.8: 垂直向量具有 $v \cdot w = 0$ 。因此 $\|v\|^2 + \|w\|^2 = \|v - w\|^2$ 。

分界线就是与 v 垂直的地方。在介于正和负之间的分界线上, $(1, -3)$ 与 $(3, 1)$ 垂直。其点积为 0。

点积反应了精确夹角 θ 。对于单位向量 u 与 U , $u \cdot U$ 的符号表明是 $\theta < 90^\circ$ 还是 $\theta > 90^\circ$ 。不仅如此, 点积 $u \cdot U$ 还是 θ 的余弦。这在 n 维中仍然正确。

单位向量 u 与 U 夹角为 θ , 有 $u \cdot U = \cos \theta$ 。无疑地, $|u \cdot U| \leq 1$ 。

记住 $\cos \theta$ 绝不会大于 1。绝不会小于 -1。单位向量的点积介于 -1 和 1 之间。 θ 的余弦由 $u \cdot U$ 显示出。

图 1.9 清晰展示了向量为 $u = (\cos \theta, \sin \theta)$ 与 $i = (1, 0)$ 时的情况。其点积为 $u \cdot i = \cos \theta$ 。这是它们夹角的余弦。

旋转任意 α 角之后, 它们依然是单位向量。向量 $i = (1, 0)$ 旋转至 $(\cos \alpha, \sin \alpha)$ 。向量 u 旋转至 $(\cos \beta, \sin \beta)$, 其中 $\beta = \alpha + \theta$ 。它们的点积为 $\cos \alpha \cos \beta + \sin \alpha \sin \beta$ 。这在三角学中是 $\cos(\beta - \alpha) = \cos \theta$ 。

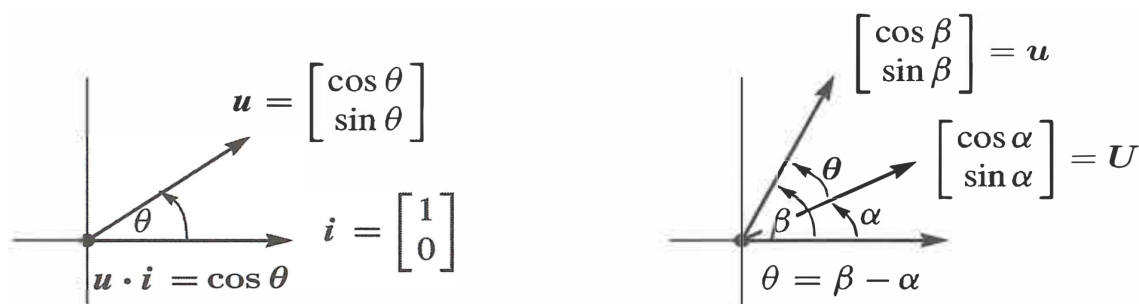


图 1.9: 单位向量: $u \cdot U$ 是 θ (夹角) 的余弦。

倘若 v 和 w 不是单位向量呢? 除以它们的长度来得出 $u = v / \|v\|$ 和 $U = w / \|w\|$ 。于是这些单位向量 u 和 U 的点积得出 $\cos \theta$ 。

余弦公式 若 v 与 w 为非零向量, 则 $\frac{v \cdot w}{\|v\| \|w\|} = \cos \theta$. (5)

不论夹角多大, 这个 $v/\|v\|$ 与 $w/\|w\|$ 的点积绝不超过 1。这就是关于点积的“施瓦茨不等式” $|v \cdot w| \leq \|v\| \|w\|$ ——或者更准确地称为柯西—施瓦茨—布尼亚科夫斯基不等式。它在法国、德国及俄罗斯（可能还有其它地方——它是数学中最重要的不等式）。

由于 $|\cos \theta|$ 一定不会超过 1, 则余弦公式给出了两个重要不等式:

施瓦茨不等式 $|v \cdot w| \leq \|v\| \|w\|$

三角不等式 $\|v + w\| \leq \|v\| + \|w\|$

例 5 对 $v = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ 与 $w = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ 求 $\cos \theta$ 并检验两个不等式。

解 其点积为 $v \cdot w = 4$ 。 v 和 w 的长度都是 $\sqrt{5}$ 。其夹角余弦是 $4/5$ 。

$$\cos \theta = \frac{v \cdot w}{\|v\| \|w\|} = \frac{4}{\sqrt{5}\sqrt{5}} = \frac{4}{5}.$$

由施瓦茨不等式, $v \cdot w = 4$ 小于 $\|v\| \|w\| = 5$ 。由三角不等式, 边 $3 = \|v + w\|$ 小于边 1+ 边 2。因为 $v + w = (3, 3)$, 则三条边为 $\sqrt{18} < \sqrt{5} + \sqrt{5}$ 。将这个三角不等式平方得出 $18 < 20$ 。

例 6 $v = (a, b)$ 与 $w = (b, a)$ 的点积为 $2ab$ 。两个向量长度都是 $\sqrt{a^2 + b^2}$ 。施瓦茨不等式 $v \cdot w \leq \|v\| \|w\|$ 说明 $2ad \leq a^2 + b^2$ 。

如果我们写成 $x = a^2$ 与 $y = b^2$, 那这就更著名了。“几何平均” \sqrt{xy} 不大于“代数平均” = 平均数 $\frac{1}{2}(x + y)$ 。

$$\text{几何平均} \leq \text{代数平均} \quad ab \leq \frac{a^2 + b^2}{2} \text{ 变为 } \sqrt{xy} \leq \frac{x + y}{2}.$$

例 5 有 $a = 2$ 和 $b = 1$ 。于是 $x = 4$ 且 $y = 1$ 。其几何平均 $\sqrt{xy} = 2$ 小于其代数平均 $\frac{1}{2}(1 + 4) = 2.5$ 。

计算说明

MATLAB, Python 及 Julia 直接处理整个向量, 而不是它们的分量。当 v 与 w 被定义出来时, 会立即获悉 $v + w$ 。按行输入 v 和 w ——' 将它们转置为列向量以做准备。 $2v + 3w$ 变成 $2 * v + 3 * w$ 。除非该行以分号结束, 否则将会打印出结果。

MATLAB $v = [2 \ 3 \ 4]'; w = [1 \ 1 \ 1]'; u = 2 * v + 3 * w$

点积 $v \cdot w$ 是一个行向量乘以一个列向量 (用 $*$ 而不是 \cdot):

我们经常见到 $\begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ 或 $v' * w$ 而不是 $\begin{bmatrix} 1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 4 \end{bmatrix}$

v 的长度被 MATLAB 称为 $\text{norm}(v)$ 。这也是 $\text{sqrt}(v' * v)$ 。然后从点积 $v' * w$ 求出余弦及具有此余弦的夹角 (弧度):

余弦公式

余弦 = $\mathbf{v}' * \mathbf{w} / (\text{norm}(\mathbf{v}) * \text{norm}(\mathbf{w}))$

反余弦

夹角 = $\text{acos}(\text{余弦})$

M 文件将创建一个新函数 `cosine(v, w)`。Python 和 Julia 都是开源的。

■ 复习关键点 ■

1. 点积 $\mathbf{v} \cdot \mathbf{w}$ 将每个分量 v_i 与 w_i 相乘然后将所有 $v_i w_i$ 加起来。
2. $\|\mathbf{v}\|$ 的长度是 $\mathbf{v} \cdot \mathbf{v}$ 的平方根。因此 $\mathbf{u} = \mathbf{v} / \|\mathbf{v}\|$ 是个单位向量：长度为 1。
3. 当向量 \mathbf{v} 与 \mathbf{w} 垂直时，其点积为 $\mathbf{v} \cdot \mathbf{w} = 0$ 。
4. θ （任意非零 \mathbf{v} 和 \mathbf{w} 的夹角）的余弦决不超过 1：

$$\text{余弦} \quad \cos \theta = \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\| \|\mathbf{w}\|} \quad \text{施瓦茨不等式} \quad |\mathbf{v} \cdot \mathbf{w}| \leq \|\mathbf{v}\| \|\mathbf{w}\|。$$