MATH1030 Inner product, norm, and orthogonality.

0. *Reminder.* Unless otherwise stated, we will deliberately confuse each (1×1)-matrix with the entry in the matrix.

1. **Definition. (Inner product in $\mathbb{R}^n$.)**

   For any $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, *write* $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^t \mathbf{v}$.

   $\langle \mathbf{u}, \mathbf{v} \rangle$ *is called the inner product of the vector* $\mathbf{u}$ *with the vector* $\mathbf{v}$.

   $\langle \cdot, \cdot \rangle$ *is called the inner product in* $\mathbb{R}^n$.

   Remark. Many people refer to $\langle \cdot, \cdot \rangle$ as the 'dot product'. A common alternative notation for $\langle \mathbf{u}, \mathbf{v} \rangle$ is $\mathbf{u} \bullet \mathbf{v}$. For this reason, we may, for convenience, read it as '$\mathbf{u}$ dot $\mathbf{v}$'.

2. **Theorem (1). (Basic properties of inner product.)**

   The statements below hold:

   (a) *Suppose* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. *Then* $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$.

   (b) *Suppose* $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ *and* $\alpha, \beta \in \mathbb{R}$. *Then* $\langle \alpha\mathbf{u} + \beta\mathbf{v}, \mathbf{w} \rangle = \alpha \langle \mathbf{u}, \mathbf{w} \rangle + \beta \langle \mathbf{v}, \mathbf{w} \rangle$.

   (c) *Suppose* $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ *and* $\alpha, \beta \in \mathbb{R}$. *Then* $\langle \mathbf{w}, \alpha\mathbf{u} + \beta\mathbf{v} \rangle = \alpha \langle \mathbf{w}, \mathbf{u} \rangle + \beta \langle \mathbf{w}, \mathbf{v} \rangle$.

   (d) *Suppose* $\mathbf{u} \in \mathbb{R}^n$. *Then* $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$. *Moreover, equality holds if and only if* $\mathbf{u} = \mathbf{0}_n$.

3. **Proof of Theorem (1).**

   (a) Suppose $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. Denote the j-th entry of $\mathbf{u}, \mathbf{v}$ as u_j, v_j respectively.
   Then $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^t \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$.
   Similarly, $\langle \mathbf{v}, \mathbf{u} \rangle = \mathbf{v}^t \mathbf{u} = v_1 u_1 + v_2 u_2 + \cdots + v_n u_n$.
   Therefore $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$.

   (b) Suppose $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ and $\alpha, \beta \in \mathbb{R}$.
   Then $\langle \alpha\mathbf{u} + \beta\mathbf{v}, \mathbf{w} \rangle = (\alpha\mathbf{u} + \beta\mathbf{v})^t \mathbf{w} = (\alpha\mathbf{u}^t + \beta\mathbf{v}^t)\mathbf{w} = \alpha\mathbf{u}^t\mathbf{w} + \beta\mathbf{v}^t\mathbf{w} = \alpha \langle \mathbf{u}, \mathbf{w} \rangle + \beta \langle \mathbf{v}, \mathbf{w} \rangle$.

   (c) Suppose $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ and $\alpha, \beta \in \mathbb{R}$.
   Then $\langle \mathbf{w}, \alpha\mathbf{u} + \beta\mathbf{v} \rangle = \langle \alpha\mathbf{u} + \beta\mathbf{v}, \mathbf{w} \rangle = \alpha \langle \mathbf{u}, \mathbf{w} \rangle + \beta \langle \mathbf{v}, \mathbf{w} \rangle = \alpha \langle \mathbf{w}, \mathbf{u} \rangle + \beta \langle \mathbf{w}, \mathbf{v} \rangle$.

   (d) Suppose $\mathbf{u} \in \mathbb{R}^n$. Denote the j-th entry of $\mathbf{u}$ as u_j respectively.
   Then $\langle \mathbf{u}, \mathbf{u} \rangle = {u_1}^2 + {u_2}^2 + \cdots + {u_n}^2 \geq 0$.
   Hence $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ if and only if $u_j = 0$ for each $j = 1, 2, \cdots, n$. The latter happens exactly when $\mathbf{u} = \mathbf{0}_n$.

4. **Definition. (Norm.)**

   For any $\mathbf{u} \in \mathbb{R}^n$, *the number* $\sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}$ *is called the norm of the vector* $\mathbf{u}$, *and is denoted by* $\|\mathbf{u}\|$.

   Remark. By definition, $\|\mathbf{u}\|^2 = \langle \mathbf{u}, \mathbf{u} \rangle$. It is often read as 'norm-square of $\mathbf{u}$'.

5. **Theorem (2). (Basic properties of norm.)**

   The statements below hold:

   (a) *Suppose* $\mathbf{u} \in \mathbb{R}^n$. *Then* $\|\mathbf{u}\| \geq 0$. *Moreover, equality holds if and only if* $\mathbf{u} = \mathbf{0}_n$.

   (b) *Suppose* $\mathbf{u} \in \mathbb{R}^n$ *and* $\alpha \in \mathbb{R}$. *Then* $\|\alpha\mathbf{u}\| = |\alpha| \cdot \|\mathbf{u}\|$.

6. **Proof of Theorem (2).**

   (a) Suppose $\mathbf{u} \in \mathbb{R}^n$.
   By definition, $\|\mathbf{u}\|^2 = \langle \mathbf{u}, \mathbf{u} \rangle \geq 0$. Then $\|\mathbf{u}\| \geq 0$.
   Moreover, $\|\mathbf{u}\| = 0$ if and only if $\langle \mathbf{u}, \mathbf{u} \rangle = 0$. The latter happens exactly when $\mathbf{u} = \mathbf{0}_n$.

   (b) Suppose $\mathbf{u} \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$.
   Then $\|\alpha\mathbf{u}\|^2 = \langle \alpha\mathbf{u}, \alpha\mathbf{u} \rangle = \alpha \cdot \alpha \langle \mathbf{u}, \mathbf{u} \rangle = \alpha^2 \langle \mathbf{u}, \mathbf{u} \rangle = \alpha^2 \|\mathbf{u}\|^2$.
   Therefore $\|\alpha\mathbf{u}\| = |\alpha| \cdot \|\mathbf{u}\|$.

7. **Theorem (3). (Conversion between inner product and norm.)**

   The statements below hold:

(a) *Suppose* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. *Then* $\|\mathbf{u} \pm \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 \pm 2\langle \mathbf{u}, \mathbf{v}\rangle$ *respectively.*

(b) *Suppose* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. *Then* $\langle \mathbf{u}, \mathbf{v}\rangle = \pm\frac{1}{2}(\|\mathbf{u} \pm \mathbf{v}\|^2 - \|\mathbf{u}\|^2 - \|\mathbf{v}\|^2)$ *respectively.*

(c) *Suppose* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. *Then* $\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2$.

Remark. The result described in Item (b) is known as the polarization identity. The result described in Item (c) is known as the parallelogramic identity.

Proof of Theorem (3). Exercise.

8. **Theorem (4). (Cauchy-Schwarz Inequality.)**

Suppose $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. *Then* $|\langle \mathbf{u}, \mathbf{v}\rangle| \leq \|\mathbf{u}\| \cdot \|\mathbf{v}\|$. *Moreover, equality holds if and only if* $\mathbf{u}, \mathbf{v}$ *are linearly dependent.*

Theorem (5). (Triangle Inequality.)

Suppose $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. *Then* $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$. *Moreover, equality holds if and only if* $\mathbf{u}, \mathbf{v}$ *are non-negative scalar multiples of each other.*

Corollary to Theorem (5). (Triangle Inequality also.)

Suppose $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. *Then* $\|\mathbf{u} - \mathbf{v}\| \geq \big|\ \|\mathbf{u}\| - \|\mathbf{v}\|\ \big|$. *Moreover, equality holds if and only if* $\mathbf{u}, \mathbf{v}$ *are non-negative scalar multiples of each other.*

9. **Definition. (Orthogonality.)**

Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. *We say* $\mathbf{u}$ *is orthogonal to* $\mathbf{v}$, *and write* $\mathbf{u} \perp \mathbf{v}$, *if and only if* $\langle \mathbf{u}, \mathbf{v}\rangle = 0$.

10. **Theorem (6). (Basic properties of orthogonality.)**

The statements below hold:

(a) *Suppose* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. *Then* $\mathbf{u} \perp \mathbf{v}$ *if and only if* $\mathbf{v} \perp \mathbf{u}$.

(b) *Suppose* $\mathbf{u} \in \mathbb{R}^n$. *Then* $\mathbf{u} \perp \mathbf{u}$ *if and only if* $\mathbf{u} = \mathbf{0}_n$.

(c) *Suppose* $\mathbf{u} \in \mathbb{R}^n$. *Then* ($\mathbf{u} \perp \mathbf{v}$ *for any* $\mathbf{v} \in \mathbb{R}^n$) *if and only if* $\mathbf{u} = \mathbf{0}_n$.

(d) *Suppose* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. *Then* $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$ *if and only if* $\mathbf{u} \perp \mathbf{v}$.

Proof of Theorem (6). Exercise (in the matrix/vector algebra).

11. **Appendix 1: Proof of the Cauchy-Schwarz Inequality.**

Suppose $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$.

(a) Suppose $\mathbf{u} = \mathbf{0}$ or $\mathbf{v} = \mathbf{0}$. Then $|\langle \mathbf{u}, \mathbf{v}\rangle| = 0 = \|\mathbf{u}\| \cdot \|\mathbf{v}\|$.

(b) From now on suppose neither of $\mathbf{u}, \mathbf{v}$ is the zero vector.
Define the quadratic polynomial $f(x)$ by $f(x) = \|\mathbf{u}\|^2 x^2 + 2\langle \mathbf{u}, \mathbf{v}\rangle x + \|\mathbf{v}\|^2$. The discriminant Δ of $f(x)$ is given by $\Delta = 4(\langle \mathbf{u}, \mathbf{v}\rangle)^2 - 4\|\mathbf{u}\|^2\|\mathbf{v}\|^2$.
For each $\alpha \in \mathbb{R}$, we have $f(\alpha) = \|\mathbf{u}\|^2\alpha^2 + 2\langle \mathbf{u}, \mathbf{v}\rangle \alpha + \|\mathbf{v}\|^2 = \cdots = \|\mathbf{u} + \alpha\mathbf{v}\|^2 \geq 0$.
Then $\Delta \leq 0$.
Therefore $4(\langle \mathbf{u}, \mathbf{v}\rangle)^2 \leq 4\|\mathbf{u}\|^2\|\mathbf{v}\|^2$.
Hence $|\langle \mathbf{u}, \mathbf{v}\rangle| \leq \|\mathbf{u}\| \cdot \|\mathbf{v}\|$.

(c) Suppose $|\langle \mathbf{u}, \mathbf{v}\rangle| = \|\mathbf{u}\| \cdot \|\mathbf{v}\|$. Then $\Delta = 0$.
Therefore the quadratic polynomial $f(x)$ has a repeated real root, say, '$x = \rho$'.
By definition $0 = f(\rho) = \|\mathbf{u}\|^2\rho^2 + 2\langle \mathbf{u}, \mathbf{v}\rangle \rho + \|\mathbf{v}\|^2 = \|\mathbf{u} - \rho\mathbf{v}\|^2$.
Then $1 \cdot \mathbf{u} - \rho\mathbf{v} = \mathbf{u} - \rho\mathbf{v} = \mathbf{0}$. Therefore $\mathbf{u}, \mathbf{v}$ are linearly dependent.

(d) Suppose $\mathbf{u}, \mathbf{v}$ are linearly dependent.
Then there exist some $\alpha, \beta \in \mathbb{R}$ such that α, β are not both zero and $\alpha\mathbf{u} + \beta\mathbf{v} = \mathbf{0}$.
Since neither of $\mathbf{u}$, $\mathbf{v}$ is the zero vector, neither of α, β is 0.
Write $\gamma = -\beta/\alpha$. We have $\mathbf{u} = \gamma\mathbf{v}$.
Then $|\langle \mathbf{u}, \mathbf{v}\rangle| = |\langle \mathbf{u}, \gamma\mathbf{u}\rangle| = |\gamma\langle \mathbf{u}, \mathbf{u}\rangle| = |\gamma\|\mathbf{u}\|^2| = |\gamma|\|\mathbf{u}\|^2 = \|\mathbf{u}\| \cdot \|\mathbf{v}\|$.

12. **Appendix 2: Proof of the Triangle Inequality.**

Suppose $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$.

(a) Suppose $\mathbf{u} = \mathbf{0}$ or $\mathbf{v} = \mathbf{0}$.

Without loss of generality, suppose $\mathbf{u} = \mathbf{0}$.

Then $\|\mathbf{u} + \mathbf{v}\| = \|\mathbf{v}\| = \|\mathbf{u}\| + \|\mathbf{v}\|$.

(b) From now on suppose neither of $\mathbf{u}, \mathbf{v}$ is the zero vector.

We have

$$\begin{aligned}(\|\mathbf{u}\| + \|\mathbf{v}\|)^2 - \|\mathbf{u} + \mathbf{v}\|^2 &= (\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\|\mathbf{u}\| \cdot \|\mathbf{v}\|) - (\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\langle \mathbf{u}, \mathbf{v}\rangle) \\ &= 2(\|\mathbf{u}\| \cdot \|\mathbf{v}\| - \langle \mathbf{u}, \mathbf{v}\rangle)\end{aligned}$$

(c) By the Cauchy-Schwarz Inequality, $\|\mathbf{u}\| \cdot \|\mathbf{v}\| \geq \langle \mathbf{u}, \mathbf{v}\rangle$.

Then $(\|\mathbf{u}\| + \|\mathbf{v}\|)^2 - \|\mathbf{u} + \mathbf{v}\|^2 \geq 0$.

Therefore $\|\mathbf{u} + \mathbf{v}\|^2 \leq (\|\mathbf{u}\| + \|\mathbf{v}\|)^2$.

Hence $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$.

(d) Suppose $\|\mathbf{u} + \mathbf{v}\| = \|\mathbf{u}\| + \|\mathbf{v}\|$.

Then $2(\|\mathbf{u}\| \cdot \|\mathbf{v}\| - \langle \mathbf{u}, \mathbf{v}\rangle) = \|\mathbf{u}\| + \|\mathbf{v}\|)^2 - \|\mathbf{u} + \mathbf{v}\|^2 = 0$.

Therefore $\|\mathbf{u}\| \cdot \|\mathbf{v}\| = \langle \mathbf{u}, \mathbf{v}\rangle$.

By the Cauchy-Schwarz Inequality, we have $|\langle \mathbf{u}, \mathbf{v}\rangle| \leq \|\mathbf{u}\| \cdot \|\mathbf{v}\| = \langle \mathbf{u}, \mathbf{v}\rangle$.

We also have $\langle \mathbf{u}, \mathbf{v}\rangle \leq |\langle \mathbf{u}, \mathbf{v}\rangle|$.

Then we have $|\langle \mathbf{u}, \mathbf{v}\rangle| = \|\mathbf{u}\| \cdot \|\mathbf{v}\| = \langle \mathbf{u}, \mathbf{v}\rangle$.

By the Cauchy-Schwarz Inequality, $\mathbf{u}, \mathbf{v}$ are linearly dependent. Then $\mathbf{u} = \gamma\mathbf{v}$ for some $\gamma \in \mathbb{R}$.

Now $|\gamma|\|\mathbf{u}\|^2 = \|\mathbf{u}\| \cdot \|\mathbf{v}\| = \langle \mathbf{u}, \mathbf{v}\rangle = \gamma\|\mathbf{u}\|^2$. Since $\mathbf{u}$ is not the zero vector, $|\gamma| = \gamma$. Then $\gamma \geq 0$.

(e) Suppose $\mathbf{u}, \mathbf{v}$ are non-negative scalar multiples of each other.

Then there is some non-negative real number ρ so that $\mathbf{v} = \rho\mathbf{u}$. Note that $1 + \rho \geq 0$.

Then $\|\mathbf{u} + \mathbf{v}\| = \|(1 + \rho)\mathbf{u}\| = |1 + \rho| \cdot \|\mathbf{u}\| = (1 + \rho)\|\mathbf{u}\| = \|\mathbf{u}\| + |\rho| \cdot \|\mathbf{u}\| = \|\mathbf{u}\| + \|\mathbf{v}\|$.

13. **Appendix 3: Geometric interpretation of the inner product, the normal, and orthogonality.**

(a) Recall the geometric interpretation of (column) vectors:

We visualize the column vector, say, $\mathbf{t} = \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{bmatrix}$ as an arrow with 'arrowhead' at the point $(t_1, t_2, \cdots, t_n)$ in real coordinate n-space and with tail at its origin.

We may then further identify this vector as the point $(t_1, t_2, \cdots, t_n)$.

By definition, $\|\mathbf{t}\| = \sqrt{{t_1}^2 + {t_2}^2 + \cdots + {t_n}^2}$.

This is the (Euclidean) distance between the origin and the point $(t_1, t_2, \cdots, t_n)$ in the coordinate n-space. Why?

- The origin, the point $(t_1, 0, 0, \cdots, 0)$ and the point with coordinate $(t_1, t_2, 0, \cdots, 0)$ are the vertices of a right-angle triangle with right angle at the point $(t_1, 0, 0, \cdots, 0)$.
  Then, by Pythagoras' Theorem, the distance between the origin and the point with $(t_1, t_2, 0, \cdots, 0)$ is given by $\sqrt{{t_1}^2 + {t_2}^2}$.
- The origin, the point $(t_1, t_2, 0, 0, \cdots, 0)$ and the point with coordinate $(t_1, t_2, t_3, 0, \cdots, 0)$ are the vertices of a right-angle triangle with right angle at the point $(t_1, t_2, 0, \cdots, 0)$.
  Then, by Pythagoras' Theorem, the distance between the origin and the point with $(t_1, t_2, t_3, 0, \cdots, 0)$ is given by $\sqrt{{t_1}^2 + {t_2}^2 + {t_3}^2}$.
- The origin, the point $(t_1, t_2, t_3, 0, 0, \cdots, 0)$ and the point with coordinate $(t_1, t_2, t_3, t_4, 0, \cdots, 0)$ are the vertices of a right-angle triangle with right angle at the point $(t_1, t_2, t_3, 0, \cdots, 0)$.
  Then, by Pythagoras' Theorem, the distance between the origin and the point with $(t_1, t_2, t_3, t_4, 0, \cdots, 0)$ is given by $\sqrt{{t_1}^2 + {t_2}^2 + {t_3}^2}$.
- So forth and so on. We deduce that the distance between the origin and the point with $(t_1, t_2, \cdots, t_{n-1}, 0)$ is given by $\sqrt{{t_1}^2 + {t_2}^2 + \cdots + {t_{n-1}}^2}$.

- The origin, the point $(t_1, t_2, \cdots, t_{n-1}, 0)$ and the point with coordinate $(t_1, t_2, \cdots, t_{n-1}, t_n)$ are the vertices of a right-angle triangle with right angle at the point $(t_1, t_2, \cdots, t_{n-1}, 0)$.
  Then, by Pythagoras' Theorem, the distance between the origin and the point with $(t_1, t_2, \cdots, t_n)$ is given by $\sqrt{{t_1}^2 + {t_2}^2 + \cdots + {t_n}^2}$.

(b) Suppose $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$. They are respectively identified as the points $(u_1, u_2, \cdots, u_n)$, $(v_1, v_2, \cdots, v_n)$ in the coordinate n-space.

Then $\|\mathbf{u} - \mathbf{v}\| = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \cdots + (u_n - v_n)^2}$ is the (Euclidean) distance between the points $(u_1, u_2, \cdots, u_n)$, $(v_1, v_2, \cdots, v_n)$ in the coordinate n-space.

(c) Suppose $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$. They are respectively identified as the points $(u_1, u_2, \cdots, u_n)$, $(v_1, v_2, \cdots, v_n)$ in the coordinate n-space.

Suppose these two points and the origin in the coordinate n-space are three non-collinear points in the coordinate n-space.

Denote by θ the angle at the origin in the triangle whose vertices are these three points.

By the Cosine Law, $\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\| \cdot \|\mathbf{v}\| \cos(\theta)$.

Then $\|\mathbf{u}\| \cdot \|\mathbf{v}\| \cos(\theta) = \frac{1}{2}(\|\mathbf{u} - \mathbf{v}\|^2 - \|\mathbf{u}\|^2 - \|\mathbf{v}\|^2) = \langle \mathbf{u}, \mathbf{v} \rangle$.

Hence $\cos(\theta) = \dfrac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\| \cdot \|\mathbf{v}\|}$.

In this context, the inequality in the result 'Cauchy-Schwarz Inequality' $| \langle \mathbf{u}, \mathbf{v} \rangle | \leq \|\mathbf{u}\| \cdot \|\mathbf{v}\|$ is simply re-stating the 'fact' that $\cos(\theta)$ has magnitude between -1 and 1.

The inequality $\|\mathbf{u}\| \leq \|\mathbf{v}\| + \|\mathbf{u} - \mathbf{v}\|$ (which is a consequence of the 'Triangle Inequality') is simply re-stating the 'fact' that the length of the line segment joining the origin to the point $(u_1, u_2, \cdots, u_n)$ is no greater than the sum of the lengths of the line segments respectively joining the origin to the point $(v_1, v_2, \cdots, v_n)$ and joining the point $(v_1, v_2, \cdots, v_n)$ to the point $(u_1, u_2, \cdots, u_n)$.

The vector $\mathbf{u}$ is orthogonal to the vector $\mathbf{v}$ exactly when the line segment joining the origin and the point $(u_1, u_2, \cdots, u_n)$ meet the line segment joining the origin with the point $(v_1, v_2, \cdots, v_n)$ at right angle.