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 $\mathbf{u} = \mathbf{0}_n$.

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 $\langle \mathbf{w}, \alpha\mathbf{u} + \beta\mathbf{v} \rangle = \alpha \langle \mathbf{w}, \mathbf{u} \rangle + \beta \langle \mathbf{w}, \mathbf{v} \rangle$.

These are collectively referred to as the bilinearity of the inner product.

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

This is referred to as the positive definite-ness of the inner product.

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

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^t\mathbf{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n.$$

Similarly,

$$\langle \mathbf{v}, \mathbf{u} \rangle = \mathbf{v}^t\mathbf{u} = v_1u_1 + v_2u_2 + \cdots + v_nu_n.$$

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 $\|\alpha\mathbf{u}\| = |\alpha| \cdot \|\mathbf{u}\|$.

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 $\|\alpha \mathbf{u}\| = |\alpha| \cdot \|\mathbf{u}\|$.

Geometric interpretation of the notion of norm.

According to Euclidean geometry, the distance from the origin of $\mathbb{R}^n$ to the 'arrow-head' of u, which is the point in $\mathbb{R}^n$ corresponding to the vector u, is exactly $\|u\|$.

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

$$\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2.$$

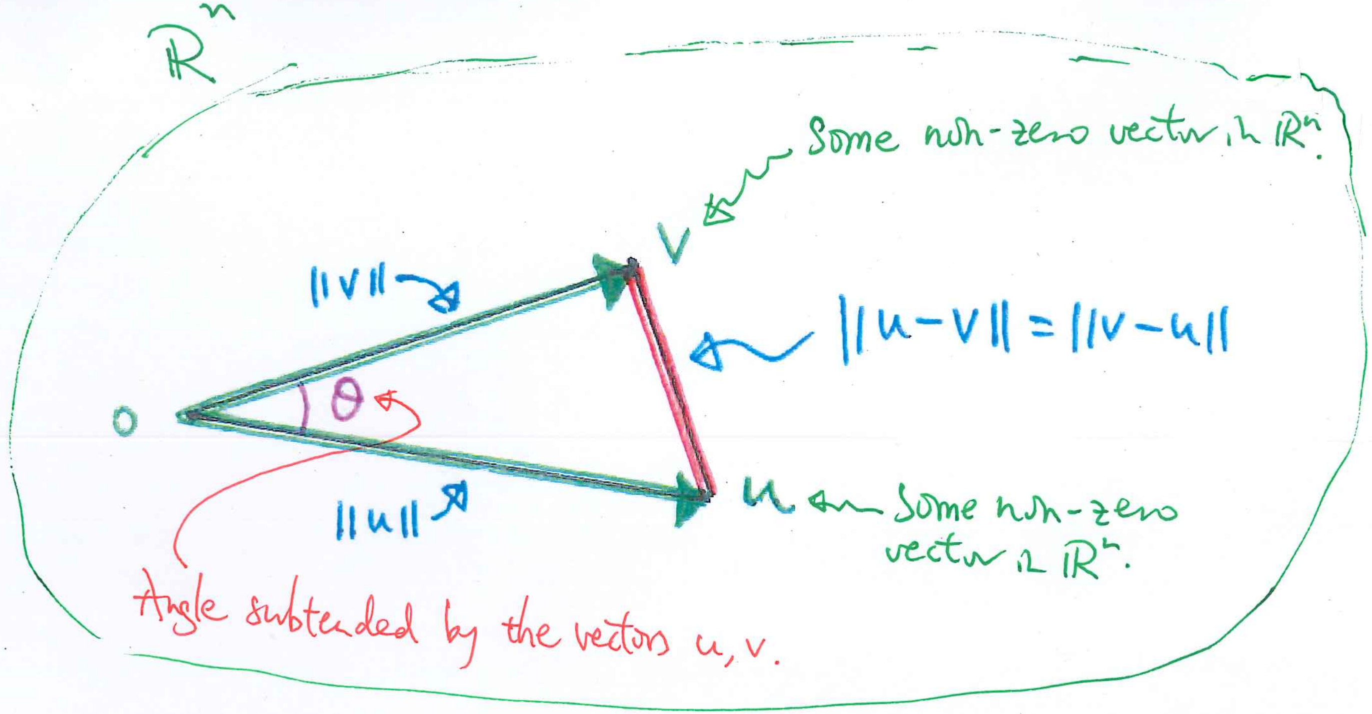
Remark.

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.

How does the inner product relate to what we learnt in elementary plane geometry and trigonometry?



Cosine Law says:

$$\|u-v\|^2 = \|u\|^2 + \|v\|^2 - 2\|u\|\cdot\|v\| \cos(\theta).$$

From what we know about inner product and norm:

$$\|u-v\|^2 = \|u\|^2 + \|v\|^2 - 2\langle u, v\rangle.$$

Then:

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

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.

How? why?

$\|u \pm v\|^2 = \langle u \pm v, u \pm v\rangle$

$= \langle u, u \pm v\rangle \pm \langle v, u \pm v\rangle$

$= \langle u, u\rangle \pm \langle u, v\rangle \pm \langle v, u\rangle + \langle v, v\rangle$

$= \langle u, u\rangle \pm 2\langle u, v\rangle + \langle v, v\rangle$

$= \|u\|^2 + \|v\|^2 \pm 2\langle u, v\rangle.$

(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.

How? why?

$\pm 2\langle u, v\rangle = \|u \pm v\|^2 - \|u\|^2 - \|v\|^2.$

$\langle u, v\rangle = \pm\frac{1}{2}\left(\|u \pm v\|^2 - \|u\|^2 - \|v\|^2\right)$

(c) *Suppose* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$.

Then

$$\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2.$$

How? why?

$$\begin{array}{rl} \|u+v\|^2 & = \|u\|^2 + \|v\|^2 + 2\langle u, v\rangle \\ +)\ \|u-v\|^2 & = \|u\|^2 + \|v\|^2 - 2\langle u, v\rangle \\ \hline \|u+v\|^2 + \|u-v\|^2 & = 2\|u\|^2 + 2\|v\|^2 \end{array}$$

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.*

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.*

Beginning steps of the argument.

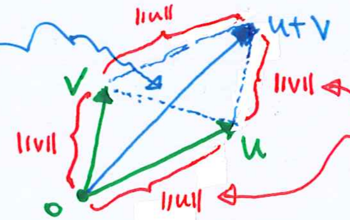
- When one of u, v is the zero vector, the result trivially holds.
- Suppose, without loss of generality, $u \neq 0$. Define the quadratic polynomial $f(t)$ by $f(t) = \|u\|^2 t^2 + 2\langle u, v\rangle t + \|v\|^2$. It happens that $f(t) = \|tu + v\|^2$. Et cetera.

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.*

How? Why? $(\|u\| + \|v\|)^2 - \|u+v\|^2 = 2(\|u\| \cdot \|v\| - \langle u, v\rangle) \geq 0$ by Cauchy-Schwarz Inequality.
Hence $\|u+v\|^2 \leq (\|u\| + \|v\|)^2$.
Then $\|u+v\| \leq \|u\| + \|v\|$. (Why?)

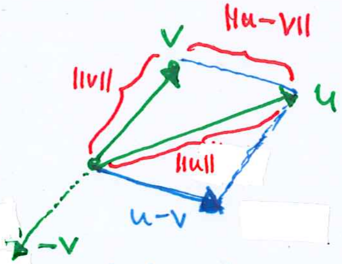
Corollary to Theorem (5). (Triangle Inequality also.)
Suppose $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$.
Then $\|\mathbf{u} - \mathbf{v}\| \geq \big|\, \|\mathbf{u}\| - \|\mathbf{v}\| \,\big|$.
Moreover, equality holds if and only if $\mathbf{u}, \mathbf{v}$ *are non-negative scalar multiples of each other.*

Geometric interpretation of Theorem (5) and its Corollary:



lengths of two sides of the triangle with vertices being the origin, the 'arrow-head' of u, and the 'arrow-head' of $u+v$.

The distance from the origin to the 'arrow-head' of $u+v$ is expected to be no less than the sum of $\|u\|$ and $\|v\|$



The sum of $\|u-v\|$ and $\|v\|$ is expected to be no less than $\|u\|$.

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).

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* $\mathbf{u} \perp \mathbf{v}$.

Pythagoras' Theorem in disguise.

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