

COMP3670/6670: Introduction to Machine Learning

Exercises with a ! denote harder ones, !! denotes very difficult, and !!! denotes optional challenge exercises.

This tutorial will be primarily about proofs in analytic geometry. There are far too many exercises to do in the 2 hours, so you should choose some particular ones to work on. Your tutor will present some in class, and feel free to post partial solutions on Piazza if you get stuck.

Question 1 Properties of the zero vector

Show that for any vector space V with any inner product $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$, we have that $\mathbf{0}$ (the zero vector) is orthogonal to every vector $\mathbf{v} \in V$.

Solution. Let \mathbf{v} be any vector in V

$$\langle \mathbf{0}, \mathbf{v} \rangle = \langle \mathbf{0} \cdot \mathbf{v}, \mathbf{v} \rangle = 0 \cdot \langle \mathbf{v}, \mathbf{v} \rangle = 0$$

Also, show that for any vector $\mathbf{v} \in V$, that $\{\mathbf{v}, \mathbf{0}\}$ forms a linearly dependant set.

Solution. Note that the equation

$$c_1 \mathbf{v} + c_2 \mathbf{0} = \mathbf{0}$$

has the non-trivial solution $c_1 = 0, c_2 = 1$, so these vectors form a linearly dependant set.

Question 2 Inner products

Prove that the standard Euclidean inner product on \mathbb{R}^2 given by

$$\mathbf{x} \cdot \mathbf{y} := x_1 y_1 + x_2 y_2$$

is an inner product.

Solution.

1. Symmetry

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 = y_1 x_1 + y_2 x_2 = \mathbf{y} \cdot \mathbf{x}$$

2. Bilinear

$$\begin{aligned} (\lambda \mathbf{x} + \phi \mathbf{y}) \cdot \mathbf{z} &= (\lambda \mathbf{x} + \phi \mathbf{y})_1 z_1 + (\lambda \mathbf{x} + \phi \mathbf{y})_2 z_2 \\ &= ((\lambda \mathbf{x})_1 + (\phi \mathbf{y})_1) z_1 + ((\lambda \mathbf{x})_2 + (\phi \mathbf{y})_2) z_2 \\ &= (\lambda x_1 + \phi y_1) z_1 + (\lambda x_2 + \phi y_2) z_2 \\ &= (\lambda x_1 + \phi y_1) z_1 + (\lambda x_2 + \phi y_2) z_2 \\ &= \lambda x_1 z_1 + \phi y_1 z_1 + \lambda x_2 z_2 + \phi y_2 z_2 \\ &= \lambda (x_1 z_1 + x_2 z_2) + \phi (y_1 z_1 + y_2 z_2) \\ &= \lambda (\mathbf{x} \cdot \mathbf{z}) + \phi (\mathbf{y} \cdot \mathbf{z}) \end{aligned}$$

We also need to show the other direction, but it follows immediately, by the property of symmetry.

3. Positive Definite

$$\mathbf{x} \cdot \mathbf{x} = x_1^2 + x_2^2 \geq 0$$

Clearly, $x_1^2 + x_2^2 = 0$ if and only if $x_1 = 0$ and $x_2 = 0$.

Question 3**Pythagorus**

We have that any inner product induces a norm,

$$\|\mathbf{x}\| := \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$$

Show that for two orthogonal vectors \mathbf{x} and \mathbf{y} (that is $\langle \mathbf{x}, \mathbf{y} \rangle = 0$) that the following holds

$$\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 = \|\mathbf{x} + \mathbf{y}\|^2$$

(This is an extension of Pythagorus' Theorem, that for a right angled triangle with hypotenuse of length c , and two other sides of length a and b , that $a^2 + b^2 = c^2$.)

Solution.

$$\|\mathbf{x} + \mathbf{y}\|^2 = \langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$$

Question 4**Parseval's Identity**

Let V be a vector space, together with an inner product $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$. Given a set of orthogonal vectors $\{x_1, \dots, x_n\}$, show that

$$\left\| \sum_{i=1}^n x_i \right\|^2 = \sum_{i=1}^n \|x_i\|^2$$

Solution.

$$\begin{aligned} \left\| \sum_{i=1}^n x_i \right\|^2 &= \left\langle \sum_{i=1}^n x_i, \sum_{i=1}^n x_i \right\rangle \\ &= \left\langle \sum_{i=1}^n x_i, \sum_{j=1}^n x_j \right\rangle \\ &= \sum_{i=1}^n \left\langle x_i, \sum_{j=1}^n x_j \right\rangle \\ &= \sum_{i=1}^n \left(\sum_{j=1}^n \langle x_i, x_j \rangle \right) \\ &= \sum_{i=1}^n \left(\langle x_i, x_i \rangle + \sum_{j \neq i} \langle x_i, x_j \rangle \right) \end{aligned}$$

Note that the property of orthogonality gives us $\langle x_i, x_j \rangle = 0$ if $i \neq j$.

$$= \sum_{i=1}^n \langle x_i, x_i \rangle = \sum_{i=1}^n \|x_i\|^2$$

Question 5**Norms**

1. Prove that the Manhattan norm (l_1 norm) on \mathbb{R}^2 defined by

$$\|\mathbf{x}\|_1 := |x_1| + |x_2|$$

is a norm. (You will need the triangle inequality on \mathbb{R} , $|a + b| \leq |a| + |b|$, to help you.)

Solution. Check the three norm axioms.

(a) Absolutely homogeneous

$$\|\lambda \mathbf{x}\|_1 = |\lambda x_1| + |\lambda x_2| = |\lambda| |x_1| + |\lambda| |x_2| = |\lambda| (|x_1| + |x_2|) = |\lambda| \|\mathbf{x}\|_1$$

(b) Positive definiteness

$$\|\mathbf{x}\|_1 = |x_1| + |x_2| \geq 0$$

Clearly $|x_1| + |x_2| = 0$ if and only if $x_1 = 0$ and $x_2 = 0$.

(c) Triangle Inequality

$$\|\mathbf{x} + \mathbf{y}\|_1 = |x_1 + y_1| + |x_2 + y_2| \leq |x_1| + |y_1| + |x_2| + |y_2| = (|x_1| + |x_2|) + (|y_1| + |y_2|) = \|\mathbf{x}\|_1 + \|\mathbf{y}\|_1$$

2. ! Prove that the supremum norm (l_∞ norm) on \mathbb{R}^2 defined by

$$\|\mathbf{x}\|_\infty := \max(|x_1|, |x_2|)$$

is a norm. (Hint: You will need triangle inequality on \mathbb{R} , and the property that if $A \subseteq B$, then $\max_{x \in A} f(x) \leq \max_{x \in B} f(x)$.)

Solution. Check the three norm axioms.

(a) Absolutely homogeneous

$$\|\lambda \mathbf{x}\|_\infty = \max\{|\lambda x_1|, |\lambda x_2|\} = \max\{|\lambda| |x_1|, |\lambda| |x_2|\} = |\lambda| \max\{|x_1|, |x_2|\} = |\lambda| \|\mathbf{x}\|_\infty$$

(b) Positive definiteness

$$\|\mathbf{x}\|_\infty = \max(|x_1|, |x_2|) \geq |x_1| \geq 0$$

Clearly $\max(|x_1|, |x_2|) = 0$ if and only if the larger of $|x_1|$ and $|x_2|$ is zero, if and only if they are both zero.

(c) Triangle Inequality

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|_\infty &= \max(|x_1 + y_1|, |x_2 + y_2|) \\ &= \max_{j \in \{1,2\}} (|x_j + y_j|) \\ &\leq \max_{i,j \in \{1,2\}} (|x_i + y_j|) \end{aligned}$$

since this allows more combinations of different components

$$\leq \max_{i,j \in \{1,2\}} (|x_i| + |y_j|)$$

by triangle inequality on \mathbb{R}

$$= \max_{i \in \{1,2\}} |x_i| + \max_{j \in \{1,2\}} |y_j|$$

as the previous term is maximised by choosing the largest value for $|x_i|$, and then largest value for $|y_j|$, and adding them together

$$= \|\mathbf{x}\|_\infty + \|\mathbf{y}\|_\infty$$

Question 6**! Basis of a vector space**

Let V be a finite dimensional vector space, and let $B = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ be a basis for V . Suppose that for any two basis vectors \mathbf{b}_i and \mathbf{b}_j , we can compute the inner product $\langle \mathbf{b}_i, \mathbf{b}_j \rangle$. Then, show that for any two vectors \mathbf{u}, \mathbf{v} in V , we can express the inner product $\langle \mathbf{u}, \mathbf{v} \rangle$ in terms of the inner product of basis vectors $\langle \mathbf{b}_i, \mathbf{b}_j \rangle$

(Hint: Use the fact that B spans the space V .)

Solution. We can express $\mathbf{u} = \sum_i u_i \mathbf{b}_i$ and $\mathbf{v} = \sum_j v_j \mathbf{b}_j$ for some collection of constants u_1, \dots, u_n and v_1, \dots, v_n , as every vector in the space can be written as some linear combination of the basis vectors. Then,

$$\begin{aligned} \langle \mathbf{u}, \mathbf{v} \rangle &= \left\langle \sum_i u_i \mathbf{b}_i, \sum_j v_j \mathbf{b}_j \right\rangle \\ &= \sum_i u_i \left\langle \mathbf{b}_i, \sum_j v_j \mathbf{b}_j \right\rangle \\ &= \sum_i u_i \sum_j v_j \langle \mathbf{b}_i, \mathbf{b}_j \rangle \\ &= \sum_{i,j} u_i v_j \langle \mathbf{b}_i, \mathbf{b}_j \rangle \end{aligned}$$

Question 7**Orthogonal matrices preserve angles and norms**

Suppose we are in the vector space \mathbb{R}^n , together with the standard Euclidean dot product, that is

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x} \cdot \mathbf{y} := \mathbf{x}^T \mathbf{y}$$

Let

$$\|\mathbf{x}\|_2 = \sqrt{\mathbf{x} \cdot \mathbf{x}}.$$

Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be an orthogonal matrix (that is, $\mathbf{A}^{-1} = \mathbf{A}^T$.)

Show that for any vector $\mathbf{x} \in \mathbb{R}^n$ that

$$\|\mathbf{Ax}\|_2 = \|\mathbf{x}\|_2$$

Using the above result (or otherwise), show that if the angle between two vectors \mathbf{x} and \mathbf{y} is θ then the angle between \mathbf{Ax} and \mathbf{Ay} is either θ , or $-\theta$ (modulo 2π).

Solution. To show the norms are the same,

$$\|\mathbf{Ax}\|_2^2 = (\mathbf{Ax})^T (\mathbf{Ax}) = \mathbf{x}^T (\mathbf{A}^T \mathbf{A}) \mathbf{x} = \mathbf{x}^T \mathbf{x} = \|\mathbf{x}\|_2^2$$

To show the angles are the same, if the angle between \mathbf{x} and \mathbf{y} is θ , then

$$\cos \theta = \frac{\mathbf{x}^T \mathbf{y}}{\|\mathbf{x}\|_2 \|\mathbf{y}\|_2}$$

Then the cosine of the angle between \mathbf{Ax} and \mathbf{Ay} is

$$\frac{(\mathbf{Ax})^T \mathbf{Ay}}{\|\mathbf{Ax}\|_2 \|\mathbf{Ay}\|_2} = \frac{\mathbf{x}^T \mathbf{A}^T \mathbf{Ay}}{\|\mathbf{x}\|_2 \|\mathbf{y}\|_2} = \frac{\mathbf{x}^T \mathbf{y}}{\|\mathbf{x}\|_2 \|\mathbf{y}\|_2}$$

which matches the above. If we let ϕ denote the angle between \mathbf{Ax} and \mathbf{Ay} , then we have that

$$\cos \theta = \cos \phi$$

Since \cos is an even function (i.e. it satisfies the property that $\cos x = \cos(-x)$) and \cos is invertible between 0 and π , either $\theta = \phi$ (which satisfies $\cos \theta = \cos \phi$) or $\theta = -\phi$.

Given an example of an orthogonal matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ and vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, such that the angle between \mathbf{x} and \mathbf{y} is not the same as the angle between \mathbf{Ax} and \mathbf{Ay} .

Solution. We choose \mathbf{A} to be the matrix that flips the plane over the x-axis, namely

$$\mathbf{A} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

It can be quickly verified that $\mathbf{A}^{-1} = \mathbf{A}^T$. Then, choose $\mathbf{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. It is clear that the angle from \mathbf{x} to \mathbf{y} is $\pi/2$. But the angle from $\mathbf{Ax} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$ to $\mathbf{Ay} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is $-\pi/2$.

Question 8

Rotation matrices preserve norms

Given a vector $\mathbf{x} \in \mathbb{R}^2$ and the rotation matrix

$$\mathbf{R}(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

Show that for any angle of rotation θ , we have

$$\|\mathbf{x}\|_2 = \|\mathbf{R}(\theta)\mathbf{x}\|_2$$

Solution.

$$\begin{aligned} \|\mathbf{R}(\theta)\mathbf{x}\|_2^2 &= \left\| \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right\|_2^2 \\ &= \left\| \begin{bmatrix} x_1 \cos \theta - x_2 \sin \theta \\ x_1 \sin \theta + x_2 \cos \theta \end{bmatrix} \right\|_2^2 \\ &= (x_1 \cos \theta - x_2 \sin \theta)^2 + (x_1 \sin \theta + x_2 \cos \theta)^2 \\ &= x_1^2 \cos^2 \theta - 2x_1 x_2 \cos \theta \sin \theta + x_2^2 \sin^2 \theta + x_1^2 \sin^2 \theta + 2x_1 x_2 \sin \theta \cos \theta + x_2^2 \cos^2 \theta \\ &= x_1^2 (\cos^2 \theta + \sin^2 \theta) + x_2^2 (\cos^2 \theta + \sin^2 \theta) \\ &= x_1^2 + x_2^2 = \|\mathbf{x}\|_2^2 \end{aligned}$$

Square rooting both sides gives the result.

Question 9

Gram-Schmidt

Let $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, the standard basis vectors for \mathbb{R}^2 . Let \mathbf{v} be any vector in \mathbb{R}^2 .

Define the projection operator

$$\text{proj}_{\mathbf{u}}(\mathbf{v}) := \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle} \mathbf{u}$$

(if $\mathbf{u} = \mathbf{0}$, then we define $\text{proj}_{\mathbf{0}}(\mathbf{v}) = \mathbf{0}$).

The Gram-Schmidt algorithm takes a set of vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ and proceeds as follows:

$$\begin{aligned} \mathbf{u}_1 &= \mathbf{v}_1 \\ \mathbf{u}_2 &= \mathbf{v}_2 - \text{proj}_{\mathbf{u}_1}(\mathbf{v}_2) \\ \mathbf{u}_3 &= \mathbf{v}_3 - \text{proj}_{\mathbf{u}_1}(\mathbf{v}_3) - \text{proj}_{\mathbf{u}_2}(\mathbf{v}_3) \\ &\dots = \dots \\ \mathbf{u}_n &= \mathbf{v}_n - \sum_{j=1}^{n-1} \text{proj}_{\mathbf{u}_j}(\mathbf{v}_n) \end{aligned}$$

The output $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ is a set of orthonormal vectors that spans the same set as $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ (If the dimension of the space spanned by the \mathbf{v}_i 's is less than n , then some of the \mathbf{u}_i 's will be zero.)

Suppose we are considering vectors in the vector space of \mathbb{R}^2 .

Show that if we input $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{v}\}$ to the Gram-Schmidt algorithm, the output is $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{0}\}$

Solution.

$$\begin{aligned}
 \mathbf{u}_1 &= \mathbf{v}_1 = \mathbf{e}_1 \\
 \mathbf{u}_2 &= \mathbf{v}_2 - \text{proj}_{\mathbf{u}_1}(\mathbf{v}_2) \\
 &= \mathbf{e}_2 - \text{proj}_{\mathbf{e}_1}(\mathbf{e}_2) \\
 &= \mathbf{e}_2 - \frac{\langle \mathbf{e}_2, \mathbf{e}_1 \rangle}{\langle \mathbf{e}_1, \mathbf{e}_1 \rangle} \mathbf{e}_1 \\
 &= \mathbf{e}_2 \\
 \text{As } \langle \mathbf{e}_1, \mathbf{e}_2 \rangle &= 0 \\
 \mathbf{u}_3 &= \mathbf{v}_3 - \frac{\langle \mathbf{v}_3, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle} \mathbf{u}_1 - \frac{\langle \mathbf{v}_3, \mathbf{u}_2 \rangle}{\langle \mathbf{u}_2, \mathbf{u}_2 \rangle} \mathbf{u}_2 \\
 &= \mathbf{v} - \frac{\langle \mathbf{v}, \mathbf{e}_1 \rangle}{\langle \mathbf{e}_1, \mathbf{e}_1 \rangle} \mathbf{e}_1 - \frac{\langle \mathbf{v}, \mathbf{e}_2 \rangle}{\langle \mathbf{e}_2, \mathbf{e}_2 \rangle} \mathbf{e}_2 \\
 &= \mathbf{v} - \langle \mathbf{v}, \mathbf{e}_1 \rangle \mathbf{e}_1 - \langle \mathbf{v}, \mathbf{e}_2 \rangle \mathbf{e}_2
 \end{aligned}$$

As \mathbf{e}_1 and \mathbf{e}_2 are the standard basis vectors for \mathbb{R}^2 , we have $\langle \mathbf{e}_1, \mathbf{e}_1 \rangle = \langle \mathbf{e}_2, \mathbf{e}_2 \rangle = 1$.

Note that we can decompose the vector \mathbf{v} as $\mathbf{v} = v_1 \mathbf{e}_1 + v_2 \mathbf{e}_2$, and we have that $\langle \mathbf{v}, \mathbf{e}_1 \rangle = v_1$ and $\langle \mathbf{v}, \mathbf{e}_2 \rangle = v_2$.

$$\begin{aligned}
 &= v_1 \mathbf{e}_1 + v_2 \mathbf{e}_2 - \langle \mathbf{v}, \mathbf{e}_1 \rangle \mathbf{e}_1 - \langle \mathbf{v}, \mathbf{e}_2 \rangle \mathbf{e}_2 \\
 &= (v_1 - \langle \mathbf{v}, \mathbf{e}_1 \rangle) \mathbf{e}_1 + (v_2 - \langle \mathbf{v}, \mathbf{e}_2 \rangle) \mathbf{e}_2 \\
 &= (v_1 - v_1) \mathbf{e}_1 + (v_2 - v_2) \mathbf{e}_2 \\
 &= \mathbf{0}
 \end{aligned}$$

Hence,

$$\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{0}\}$$

as required.

Question 10

!!! Cauchy-Schwartz

Prove the Cauchy-Schwartz inequality for a general inner product and corresponding induced norm:

$$\langle \mathbf{u}, \mathbf{v} \rangle \leq \|\mathbf{u}\| \|\mathbf{v}\|$$

(Hint: Let $\mathbf{z} = \mathbf{u} - \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \mathbf{v}$, and start with the fact that $\langle \mathbf{z}, \mathbf{z} \rangle \geq 0$.)

Solution. As the hint suggests, let $\mathbf{z} = \mathbf{u} - \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \mathbf{v}$.

$$\begin{aligned}
 \langle \mathbf{z}, \mathbf{z} \rangle &\geq 0 \\
 \left\langle \mathbf{u} - \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \mathbf{v}, \mathbf{u} - \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \mathbf{v} \right\rangle &\geq 0 \\
 \langle \mathbf{u}, \mathbf{u} \rangle - \left\langle \mathbf{u}, \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \mathbf{v} \right\rangle - \left\langle \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \mathbf{v}, \mathbf{u} \right\rangle + \left\langle \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \mathbf{v}, \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \mathbf{v} \right\rangle &\geq 0 \\
 \|\mathbf{u}\|^2 - \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \langle \mathbf{u}, \mathbf{v} \rangle - \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \langle \mathbf{v}, \mathbf{u} \rangle + \frac{\langle \mathbf{u}, \mathbf{v} \rangle^2}{\langle \mathbf{v}, \mathbf{v} \rangle^2} \langle \mathbf{v}, \mathbf{v} \rangle &\geq 0 \\
 \|\mathbf{u}\|^2 - 2 \frac{\langle \mathbf{u}, \mathbf{v} \rangle^2}{\langle \mathbf{v}, \mathbf{v} \rangle} + \frac{\langle \mathbf{u}, \mathbf{v} \rangle^2}{\langle \mathbf{v}, \mathbf{v} \rangle^2} \langle \mathbf{v}, \mathbf{v} \rangle &\geq 0
 \end{aligned}$$

as $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$

$$\begin{aligned}\|\mathbf{u}\|^2 - 2\frac{\langle \mathbf{u}, \mathbf{v} \rangle^2}{\langle \mathbf{v}, \mathbf{v} \rangle} + \frac{\langle \mathbf{u}, \mathbf{v} \rangle^2}{\langle \mathbf{v}, \mathbf{v} \rangle} &\geq 0 \\ \|\mathbf{u}\|^2 - \frac{\langle \mathbf{u}, \mathbf{v} \rangle^2}{\|\mathbf{v}\|^2} &\geq 0 \\ \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 - \langle \mathbf{u}, \mathbf{v} \rangle^2 &\geq 0 \\ \langle \mathbf{u}, \mathbf{v} \rangle &\leq \|\mathbf{u}\| \|\mathbf{v}\|\end{aligned}$$

as required.