Inner Product Spaces (Part I)

Conceptual knowledge

- Revise fundamental concepts of vectors and extend these ideas to higher dimensions/vector spaces.
- Understand the relationship between dot product and inner product, and what an inner product space is.
- Familiarity with the Cauchy-Schwarz and triangle inequalities.
- Understand orthogonality and the Pythagorean theorem.

Procedural knowledge

- Find magnitude, dot products, unit vectors, distance/angle between vectors, inner products.
- Determine inner product spaces/calculating inner products.
- Find orthogonal projections.

1 Vector Fundamentals

This section will recap some basic vector ideas in \mathbb{R}^2 and update our notation & understanding so that we can extend them to higher dimensions.

1.1 Norm

We define the \underline{norm} (also called \underline{length} or $\underline{magnitude}$) of a vector to be,

$$\| \underline{\vee} \| = \overline{\int V_1^2 + V_2^2} , \underline{\vee} = (\underline{\vee}_1, \underline{\vee}_1) \in \mathbb{R}^2$$

Notice we are now using double lines to denote this quantity.

The single line version refers to absolute value, which is a particular type of norm (more on this later).

This definition can be analogously extended to the n^{th} dimension as,

$$\underline{V} = (\underline{V}_{11}, \dots, V_n) \in \mathbb{R}^n \implies ||\underline{V}|| = \overline{\int V_1^2 + \dots + V_n^2}$$

This definition gives us an idea of the total size of the vector in relation to its components, irrespective of direction (no minus signs), and also reduces down to the well known Pythagorean theorem in 2 and 3 dimensions.

It can also be easily shown that dividing any vector by its own norm results in a vector which has a norm of 1. This is called a <u>unit vector</u>, and it points in the direction of the original vector.

This is often called <u>"normalising"</u> the vector, and in general normalising with respect to some fixed value is done all the time in mathematics in order to make relationships more clear.

Note that for any $c \in \mathbb{R}$ we have,

The law of cosines gives,

1.2 Distance & Angle Between Vectors

We define the <u>distance</u> between two vectors, **u** and **v**, to be the <u>norm</u> of the resultant difference vector:

$$d(\underline{U},\underline{Y}) = ||\underline{U} - \underline{Y}|| = \int (\underline{U}_1 - \underline{V}_1)^2 + \dots + (\underline{U}_n - \underline{V}_n)^2$$

$$||\underline{Y}|| = ||\underline{U} - \underline{Y}|| = \int (\underline{U}_1 - \underline{V}_1)^2 + \dots + (\underline{U}_n - \underline{V}_n)^2$$

$$||\underline{Y}|| = ||\underline{U} - \underline{Y}|| = \int (\underline{U}_1 - \underline{V}_1)^2 + \dots + (\underline{U}_n - \underline{V}_n)^2$$

$$||\underline{Y}|| = \int (\underline{U}_1 - \underline{V}_1)^2 + \dots + (\underline{U}_n - \underline{V}_n)^2$$

$$||\underline{Y}|| = \int (\underline{U}_1 - \underline{V}_1)^2 + \dots + (\underline{U}_n - \underline{V}_n)^2$$

$$||\underline{Y}|| = \int (\underline{U}_1 - \underline{V}_1)^2 + \dots + (\underline{U}_n - \underline{V}_n)^2$$

$$||\underline{Y}|| = \int (\underline{U}_1 - \underline{V}_1)^2 + \dots + (\underline{U}_n - \underline{V}_n)^2$$

$$||\underline{Y}|| = \int (\underline{U}_1 - \underline{V}_1)^2 + \dots + (\underline{U}_n - \underline{V}_n)^2$$

$$||\underline{Y}|| = \int (\underline{U}_1 - \underline{V}_1)^2 + \dots + (\underline{U}_n - \underline{V}_n)^2$$

$$||\underline{Y}|| = \int (\underline{U}_1 - \underline{V}_1)^2 + \dots + (\underline{U}_n - \underline{V}_n)^2$$

$$||\underline{Y}|| = \int (\underline{U}_1 - \underline{V}_1)^2 + \dots + (\underline{U}_n - \underline{V}_n)^2$$

The numerator is the familiar dot product (or scalar product).

We generalise the dot product definition as,

We would like to write that the cosine of the angle between 2 n-dimensional vectors is their dot product divided by their norms multiplied together, however we can only measure angles in 3 dimensions at most.

So the meaning of the angle is ambiguous in n > 3 dimensions and our equation would only hold if we can guarantee that our vector equation is between -1 and 1. We can prove this with the Cauchy-Schwarz inequality, but first let's look at some properties of the dot product.

Dot Product Properties

3.
$$C(\overline{\Lambda},\overline{\Lambda}) = \overline{\Lambda}, C\overline{\Lambda} = C\overline{\Lambda}, \overline{\Lambda}$$

3. $C(\overline{\Lambda},\overline{\Lambda}) = \overline{\Lambda}, C\overline{\Lambda} = C\overline{\Lambda}, \overline{\Lambda}$

4.
$$\forall \cdot \forall = 1 \forall 1$$

These properties are easy enough to prove using the definitions, but are necessary for the generalisation to inner product which we will be looking at shortly.

Defining vectors in \mathbb{R}^n with vector addition, scalar multiplication, norm and dot product, we call the resulting vector space the *Euclidean n-space*.

The Cauchy-Schwarz Inequality 1.3

Given 2 vectors, \mathbf{u} and \mathbf{v} , in \mathbb{R}^n then,

Note that the inequality can also be written as,
$$\left(\sum_{i=1}^{n} v_{i} v_{i}^{i} \right)^{2} = \sum_{i=1}^{n} v_{i}^{2} \sum_{i=1}^{n} v_{i}^{2}$$

$$3$$

Proof Consider,
$$\sum_{i=1}^{n} \sum_{j=1}^{n} (v_i v_j - v_j v_i)^2 > 0$$

$$\Rightarrow \sum_{i=1}^{n} \sum_{j=1}^{n} v_i^2 v_j^n + v_j^2 v_i^2 - 2v_i v_j v_j v_i > 0$$

$$\sum_{i=1}^{n} \sum_{j=1}^{n} v_i^2 + \sum_{j=1}^{n} v_j^2 \sum_{i=1}^{n} v_i^2 - 2\sum_{i=1}^{n} v_i v_i \geq 0$$

$$\sum_{i=1}^{n} v_i^2 \sum_{j=1}^{n} v_i^2 - 2\left(\sum_{i=1}^{n} v_i v_i\right)^2 > 0$$

$$\sum_{i=1}^{n} v_i^2 \sum_{i=1}^{n} v_i^2 - 2\left(\sum_{i=1}^{n} v_i v_i\right)^2 \leq \sum_{i=1}^{n} v_i^2 \sum_{i=1}^{n} v_i^2$$

$$\sum_{i=1}^{n} v_i v_i = \sum_{i=1}^{n} v_i = \sum_{i=1}^{n} v_i v_i = \sum_$$

We then have,

which guarantees it is in the range of cosine and therefore we define the angle between non-zero *n*-dimensional vectors to be,

$$\frac{||Y||||X||}{||X||||X||} = \cos \theta, \quad 0 \leq \theta \leq \pi$$

Example 1.1'

Find the angle between the 4th dimensional vectors, $\mathbf{u} = (-4, 0, 2, -2)$ and $\mathbf{v} = (2, 0, -1, 1)$.

$$\cos \phi = -8 + 0 - 2 - 2 = -12 = -1$$

$$\int 2\sqrt{6} = -12 = -1$$

$$2\sqrt{6} = -12 = -1$$

$$1 = -1$$

Definition

Vectors **u** and **v** in \mathbb{R}^n are **orthogonal** if,

$$U \cdot V = 0$$

1.4 Triangle Inequality

This inequality has many applications across mathematics so you should remember it.

We will prove it using the Cauchy-Schwarz inequality.

Given 2 vectors,
$$\mathbf{u}$$
 and \mathbf{v} , in \mathbb{R}^n then,
$$\| \underline{\mathcal{U}} + \underline{\mathcal{V}} \| \leq \| \underline{\mathcal{U}} \| + \| \underline{\mathcal{V}} \|$$

$$\frac{\|U + Y\|^{2}}{\|U + Y\|^{2}} = (U + Y) \cdot (U + Y)$$

$$= U \cdot (U + Y) + Y \cdot (U + Y)$$

$$= \|U \|^{2} + \|Y \|^{2} + 2 U \cdot Y$$

$$\leq \|U \|^{2} + \|Y \|^{2} + 2 \|U \| \|Y \|$$

$$\leq \|U \| + \|Y \|^{2} + 2 \|U \| \|Y \|$$

$$\leq (\|U \| + \|Y \|)^{2}$$

$$= D$$

Note that this can be seen as a generalisation of Pythagoras' theorem in n dimensions (consider the case where \mathbf{u} and \mathbf{v} are orthogonal).

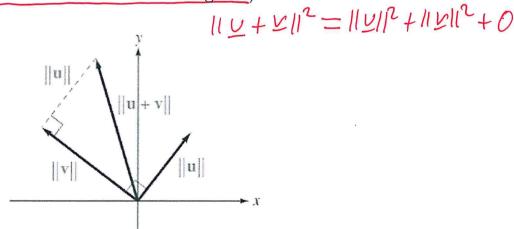


Figure 1: Triangle inequality in 2 dimensions © Pearson, 2016.

2 Inner Product Space Fundamentals

Having generalised 2D/3D concepts of length, distance and angle to the n^{th} dimension of the reals, \mathbb{R}^n , we now generalise these concepts to vector spaces.

The dot product we have seen in \mathbb{R}^n is known as the <u>Euclidean inner product</u>, and is a type of inner product.

Inner products are just functions which satisfy the following 4 axioms.

Definition

Let \mathbf{u} , \mathbf{v} and \mathbf{w} be in a vector space, V, and c be a scalar. An <u>inner product</u> is a function that associates a real number, $\langle \mathbf{u}, \mathbf{v} \rangle$, with the 2 vectors such that,

$$1. \angle U, \underline{V} > = \angle \underline{V}, \underline{V} >$$

Definition

An <u>inner product space</u> is a vector space which has an inner product defined over it.

We can easily show that the axioms hold for the Euclidean inner product (dot product in \mathbb{R}^n).

Let us look at a non-standard inner product on \mathbb{R}^2 .

Example 2.1

Show that the following function defines an inner product on \mathbb{R}^2 .

$$\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1 + 2u_2 v_2$$
 be finition

1)
$$\angle U, \underline{V} >= U, \underline{V}_1 + 2u_2V_2 = \underline{V}_1 + 2v_2v_2 = \underline{Z}_2 \underline{V}_2 >$$
2) $\angle U, \underline{V} + \underline{W} > = U_1(v_1 + w_1) + 2u_2(v_2 + w_2)$

$$= (U_1 V_1 + 2u_2V_2) + (u_1w_1 + 2u_2w_2)$$

$$= \angle U, \underline{V} > + \angle U, \underline{W} >$$
3) $C \angle U, \underline{V} > = C(u_1v_1 + 2u_2v_2)$

$$= (cu_1)v_1 + 2(cu_2)v_2 = \underline{Z}_2 \underline{V}_2 >$$

$$+ (\underline{U}, \underline{V}) = \underline{U}_1 \underline{U}_1 + 2u_2v_2 = \underline{U}_1^2 + 2u_2^2 > 0$$
Also, $\underline{U}_1^2 + 2u_2^2 = 0 = \underline{V}_1 = u_2 = 0$ is only real solution

Example 2.2

Show that the following function is not an inner product on \mathbb{R}^3 .

$$\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1 - 2u_2 v_2 + u_3 v_3$$

4) Let
$$U = (0,1,0) \Longrightarrow \angle U, U \succcurlyeq 0 - 2 + 0 = -2 \angle 0$$

So the question is why do we need these different inner products?

There are applications all across mathematics, but within the context of linear algebra, they can be used to construct methods of solving extremely large systems of equations that are more efficient and robust than the methods we have seen so far for the relatively small systems we have been dealing with.

For more information you can look into Krylov subspace methods.

The next example will be familiar to you from advanced calculus.

Example 2.3

Let f and g be real-valued continuous functions in the vector space, C[a,b] (the set of continuous functions over the interval [a,b]). Show that the following defines an inner product on C[a,b].

$$\langle f, g \rangle = \int_a^b f(x)g(x) dx$$

1)
$$\angle y, y > \Rightarrow \angle f, g > = \int_a^b f(x)g(x)dx = \int_a^b g(x)f(x)dx = \angle g, f$$

2) $\angle f, g + h > = \int_a^b f(x)[g(x) + h(x)]dx = \int_a^b f(x)g(x) + f(x)h(x)dx$

$$= \int_a^b f(x)g(x)h + \int_a^b f(x)h(x)dx = \angle f, g >$$

$$= \int_a^b f(x)g(x)h + \int_a^b f(x)h(x)dx = \angle f, g >$$

3) $c \angle f, g > = c \int_a^b f(x)g(x)dx = \int_a^b [cf(x)]g(x)dx = \angle cf, g >$
4) $\angle f, f > = \int_a^b [f(x)]^2 dx = 0 \Rightarrow f(x) = 0$ (Since all are general)

The previous inner product, along with the definition of orthogonality are used to derive Fourier series.

2.1 Definitions on Inner Product Spaces

We now generalise the concepts of vectors in \mathbb{R}^n to general vector spaces.

Even though the geometric meaning is lost, the definitions form a natural extension and when reduced down to the lower dimensions we recover the geometric qualities.

Given \mathbf{u} and \mathbf{v} in an inner product space, V, we have the following definitions.

1. The norm (or length) of **u** is,
$$\parallel \underline{U} \parallel = \int \underline{L}\underline{U}, \underline{U} >$$

2. The distance between **u** and **v** is,
$$d(\underline{v},\underline{v}) = 11 \ \underline{V} - \underline{v} \ 1$$

3. The angle between
$$\mathbf{u} \neq 0$$
 and $\mathbf{v} \neq 0$ is,
$$\cos \theta = \underbrace{\angle \cup, \vee}_{\parallel \cup \parallel \parallel \parallel \vee \parallel}$$

4. The vectors **u** and **v** are orthogonal if,
$$\angle \underline{\cup}, \underline{\vee} > = 0$$

We still have that a unit vector is when its norm is equal to 1.

Note that orthogonality depends on the inner product used. Vectors which are orthogonal with respect to one inner product are often not orthogonal with respect to another.

Example 2.4

Given polynomials, p and q, in the vector space, P_n (set of polynomial of degree n or less), define an inner product as,

$$\langle p, q \rangle = \sum_{i=0}^{n} a_i b_i$$

$$\angle P, q > = | \times 4 + 0 \times -2 + -2 \times 1 = 2$$

 $\angle q, r > = 4 \times 0 + -2 \times 1 + 1 \times 2 = 0$ orthogonal
 $||q|| = \int \angle q, q > = \int 4^2 + (-2)^2 + 1^2 = \int 21$
 $||q|| = ||q|| = ||(1 - 2 \times 2) - (4 - 2 \times + 2)|| = ||-3 + 2 \times -3 \times 2||$
 $= \int q + 4 + q = \int 22$

Example 2.5

Use the inner product from Example 2.3, with f(x) = x and $g(x) = x^2$ on C[0, 1] to find ||f|| and d(f, g).

$$\|f\|^{2} = \langle f, f \rangle = \int_{0}^{\infty} (a)(x) dx = \int_{0}^{\infty} x^{2} dx = \left[\frac{x^{3}}{3} \right]_{0}^{\infty} = \frac{1}{3}$$

$$\|f\|^{2} = \langle f, f \rangle = \int_{0}^{\infty} (a)(x) dx = \int_{0}^{\infty} x^{2} dx = \left[\frac{x^{3}}{3} \right]_{0}^{\infty} = \frac{1}{3}$$

$$\|f\|^{2} = \langle f, f \rangle = \int_{0}^{\infty} (a)(x) dx = \int_{0}^{\infty} x^{2} dx = \left[\frac{x^{3}}{3} \right]_{0}^{\infty} = \frac{1}{3}$$

$$\|f\|^{2} = \langle f, f \rangle = \int_{0}^{\infty} (a)(x) dx = \int_{0}^{\infty} x^{2} dx = \left[\frac{x^{3}}{3} \right]_{0}^{\infty} = \frac{1}{3}$$

$$\|f\|^{2} = \langle f, f \rangle = \int_{0}^{\infty} (a)(x) dx = \int_{0}^{\infty} x^{2} dx = \left[\frac{x^{3}}{3} \right]_{0}^{\infty} = \frac{1}{3}$$

$$\|f\|^{2} = \langle f, f \rangle = \int_{0}^{\infty} (a)(x) dx = \int_{0}^{\infty} x^{2} dx = \left[\frac{x^{3}}{3} \right]_{0}^{\infty} = \frac{1}{3}$$

$$\|f\|^{2} = \langle f, f \rangle = \int_{0}^{\infty} (a)(x) dx = \int_{0}^{\infty} x^{2} dx = \left[\frac{x^{3}}{3} \right]_{0}^{\infty} = \frac{1}{3}$$

$$\|f\|^{2} = \int_{0}^{\infty} (a)(x) dx = \int_{0}^{\infty} x^{2} dx = \left[\frac{x^{3}}{3} \right]_{0}^{\infty} = \frac{1}{3}$$

$$\|f\|^{2} = \int_{0}^{\infty} (a)(x) dx = \int_{0}^{\infty} x^{2} dx = \left[\frac{x^{3}}{3} \right]_{0}^{\infty} = \frac{1}{3}$$

$$\|f\|^{2} = \int_{0}^{\infty} (a)(x) dx = \int_{0}^{\infty} x^{2} dx = \left[\frac{x^{3}}{3} \right]_{0}^{\infty} = \frac{1}{3}$$

$$\|f\|^{2} = \int_{0}^{\infty} (a)(x) dx = \int_{0}^{\infty} x^{2} dx = \left[\frac{x^{3}}{3} \right]_{0}^{\infty} = \frac{1}{3}$$

$$\|f\|^{2} = \int_{0}^{\infty} (a)(x) dx = \int_{0}^{\infty} x^{2} dx = \left[\frac{x^{3}}{3} \right]_{0}^{\infty} = \frac{1}{3}$$

$$\|f\|^{2} = \int_{0}^{\infty} (a)(x) dx = \int_{0}^{\infty} x^{2} dx = \left[\frac{x^{3}}{3} \right]_{0}^{\infty} = \frac{1}{3}$$

$$\|f\|^{2} = \int_{0}^{\infty} (a)(x) dx = \int_{0}^{\infty} x^{2} dx = \left[\frac{x^{3}}{3} \right]_{0}^{\infty} = \frac{1}{3}$$

$$\|f\|^{2} = \int_{0}^{\infty} (a)(x) dx = \int_{0}^{\infty} x^{2} dx = \left[\frac{x^{3}}{3} \right]_{0}^{\infty} = \frac{1}{3}$$

$$\|f\|^{2} = \int_{0}^{\infty} (a)(x) dx = \int_{0}^{\infty} x^{2} dx = \int_{0}^{\infty} x^{3} dx = \int_{$$

$$= \left[\frac{3}{3} - \frac{4}{5} + \frac{5}{5} \right]_{0}^{1} = \frac{1}{30}$$

$$\frac{1}{30} = \frac{1}{30}$$

We restate the Cauchy-Schwarz inequality, generalised to inner products:

The triangle inequality and Pythagorean theorem also still apply.

Example 2.6

Verify the Cauchy-Schwarz inequality for the inner product defined in Example 2.3, with f(x) = x and $g(x) = \cos x$ on $C[0, \pi]$.

2.2 Orthogonal Projections

We defined projections of one vector onto another in 2 (and 3) dimensions as shown in the following figure.

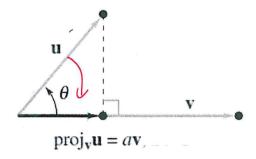


Figure 2: Projection of vector **u** onto **v** © Pearson, 2016.

Geometrically we can think of the shadow of the vector, \mathbf{u} , onto the vector, \mathbf{v} .

It is simply a scalar multiplied by \mathbf{v} , such that,

We then have,
$$\alpha = \frac{|U||V||}{|V||^2}$$

$$\alpha = \frac{|U||V||}{|V||^2}$$
We then have,
$$\alpha = \frac{|U||V||}{|V||^2}$$

Example 2.7

Find the orthogonal projection of $\mathbf{u} = (4, 2)$ onto $\mathbf{v} = (3, 4)$.

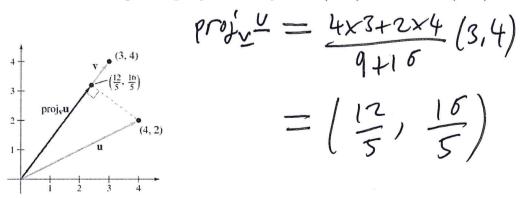


Figure 3: Figure for Example 2.7 © Pearson, 2016.

The extension of vector projection for inner product spaces is given below.

Definition

Let **u** and **v** be vector in an inner product space, V, with $\mathbf{v} \neq \mathbf{0}$. The **orthogonal projection** of **u** onto **v** is given by,

Example 2.8

Pot product

Find the orthogonal projection of $\mathbf{u} = (6, 2, 4)$ onto $\mathbf{v} = (1, 2, 0)$ with the Euclidean inner product in \mathbb{R}^3 .

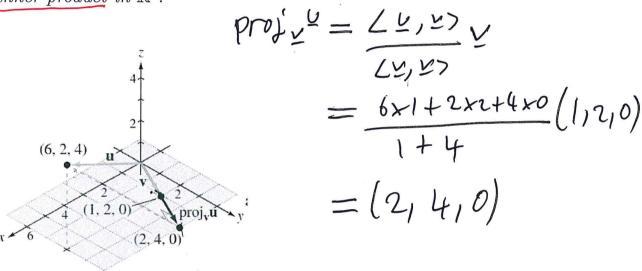


Figure 4: Figure for Example 2.8 © Pearson, 2016.

Note that in \mathbb{R}^3 the line which we project down onto \mathbf{v} lies in the plane spanned by \mathbf{u} and \mathbf{v} .

The line we project down is $\mathbf{u} - \text{proj}_{\mathbf{v}}\mathbf{u}$.

Let **u** and **v** be vector in an inner product space, V, with $\mathbf{v} \neq \mathbf{0}$. Then,

The theorem says that the line, $\mathbf{u} - \text{proj}_{\mathbf{v}} \mathbf{u}$, is the shortest distance between \mathbf{u} and a point on \mathbf{v} .

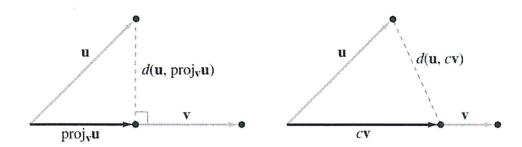


Figure 5: Illustration of the projection distance theorem © Pearson, 2016.

Proof

RHS;
$$d(y,cy) = \|y-cy\|$$
, Let $a = \angle y, y > 2$
 $\|y-cy\|^2 = \|y-ay+ay-cy\|^2$
 $= \|(y-ay)+(a-c)y\|^2$

But $(y-ay)$ and $(a-c)y$ are orthogonal;

From the triangle inequality for orthogonal vectors; $||(V-aV)+(a-c)V||^2=||V-aV||^2+||(a-c)V||^2$

$$\| \underline{U} - C\underline{V} \|^2 = \| \underline{U} - \underline{a} \underline{V} \|^2 + \| (\underline{a} - c) \underline{V} \|^2$$
Since $\underline{a} \neq \underline{c}$ and $\underline{V} \neq 0$, $(\underline{a} - c)^2 \| \underline{V} \|^2 > 0$

Since
$$av = \text{proj}_{\underline{v}} \underline{v}$$
 we have,
$$d(\underline{v}, \text{proj}_{\underline{v}} \underline{v}) \angle d(\underline{v}, \underline{c}\underline{v})$$

$$for c \neq \underline{\angle v}, \underline{v} > \underline{U}$$

Example 2.9

Let f(x) = 1 and g(x) = x be function in C[0,1] with the inner product defined as in Example 2.3. Find the orthogonal projection of f onto g.

in Example 2.3. Find the orthogonal projection of
$$f$$
 onto g .

Let f and f be a sum of f onto g .

Let f and f be a sum of f onto g .

Let f and f be a sum of f onto g .

Let f and f be a sum of f onto g .

Let f and f be a sum of f onto g .

Let f and f be a sum of f onto g .

Let f and f be a sum of f onto g .

Let f and f be a sum of f onto g .

Let f and f be a sum of f onto g .

Let f and f be a sum of f onto g .

Let f and f be a sum of f onto g .

Let f and f be a sum of f onto g .

Let f and f be a sum of f onto g .

Let f and f be a sum of f onto g .

Let f and f be a sum of f onto g .

Let f and f be a sum of f onto g .

Let f and f be a sum of f be a sum of f onto g .

Let f and f be a sum of f b