Mathematics for Economics: Linear Algebra **Anthony Tay**

The Dot Product

The Linear Combination

We will be dealing with expressions such as

$$a_1x_1 + a_2x_2 + ... + a_kx_k$$

a lot. We call this a linear combination of the items $x_1, x_2, ..., x_k$.

E.g. Suppose you buy $n_1, n_2, ..., n_k$ number of k goods priced $p_1, p_2, ..., p_k$ respectively. Your total expenditure is $n_1p_1 + n_2p_2 + ... + n_kp_k$.

The dot product is a central concept for such type of computations. Define a k-dimensional vector to be an ordered list of k numbers, e.g.,

$$\mathbf{a} = (a_1, a_2, ..., a_k).$$

Ordered means that the position is important. The vectors (1,2,3,4) is different from (2,1,3,4) even though the vectors contain the same numbers. Given two vectors

$$\mathbf{a} = (a_1, a_2, ..., a_k) \text{ and } \mathbf{x} = (x_1, x_2, ..., x_k),$$

define their dot product to be

$$\mathbf{a} \cdot \mathbf{b} = a_1x_1 + a_2x_2 + ... + a_kx_k.$$

Sometimes the dot product is also called the 'inner product'.

We want to learn how to manipulate linear combinations. This means learning how to manipulate the dot product. However, before listing the "rules for the dot product", we make two more very important definitions. Given a vector $\mathbf{x} = (x_1, x_2, ..., x_k)$, the scalar multiple of $\mathbf{x}$ and a number α is defined to be the vector

$$\alpha\mathbf{x} = (\alpha x_1, \alpha x_2, ..., \alpha x_k).$$

Given two vectors of the same dimension

$$\mathbf{x} = (x_1, x_2, ..., x_k) \qquad \text{and} \qquad \mathbf{y} = (y_1, y_2, ..., y_k)$$

their sum is defined to be

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, ..., x_k + y_k).$$

Obviously, to be able to add two vectors, they must be of the same dimension.

Theorem (Properties of the dot product) Let $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$ be arbitrary k-dimensional vectors. Then

(i) $\mathbf{x} \cdot \mathbf{x} \geq 0$ with $\mathbf{x} \cdot \mathbf{x} = 0$ if and only if $\mathbf{x}$ is the zero vector $\mathbf{0} = (0, 0, ..., 0)$.

(ii) $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$

(iii) $(\mathbf{x} + \mathbf{y}) \cdot \mathbf{z} = \mathbf{x} \cdot \mathbf{z} + \mathbf{y} \cdot \mathbf{z}$

(iv) $(\alpha\mathbf{x}) \cdot \mathbf{y} = \alpha(\mathbf{x} \cdot \mathbf{y})$

Proof: Exercise!

Some Two-Dimensional Geometry

We now talk specifically about two-dimensional vectors, and give them a geometric interpretation. Given a two-dimensional vector $\mathbf{u} = (u_1, u_2)$, we can view it as a point in two-dimensional space (take the first number to represent to horizontal "x" coordinate, and the second number to represent the vertical "y" coordinate). You can also visualize it as an arrow from the origin to the point (u_1, u_2).

You know that the sum of two vectors $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$ gives you the vector diagonal of the parallelogram defined by the two vectors. You can also view it as though you either (i) moved the origin to the point (u_1, u_2) then moved the point (u_1, u_2) in the direction and distance defined by the vector $\mathbf{v}$ to the point $(u_1 + v_1, u_2 + v_2)$, or (ii) moved the origin to the point (v_1, v_2), then moved that point in the direction and distance defined by $\mathbf{u}$, to the point $(u_1 + v_1, u_2 + v_2)$. This tells you three things. First, $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$. Second, the essential features of a vector are its direction and length. Third, the starting point is not important. Apart from the direction and length, the position of the vector $\mathbf{u} = (u_1, u_2)$ is not important. The vector $\mathbf{u}$ can be viewed as moving the origin u_1 units in the horizontal direction, and u_2 units vertically. The vector $\mathbf{v} + \mathbf{u}$ can be viewed as moving the point $\mathbf{v}$ u_1 units in the horizontal direction, and u_2 units vertically. The arrows representing $\mathbf{u}$ either starting from the origin, or starting from $\mathbf{v}$ can be considered the same vector.

Exercise: let $\mathbf{u} = (4, 2)$ and $\mathbf{v} = (-1, 2)$. Illustrate $\mathbf{u} + \mathbf{v}$ and $\mathbf{u} - \mathbf{v}$.

A scalar multiple of a vector $\alpha\mathbf{u}$ stretches the vector by a factor of α. If α is negative, the direction is opposite to the original vector.

Exercise: let $\mathbf{u} = (4, 2)$ and $\mathbf{v} = (-1, 2)$. Illustrate $2\mathbf{u}$ and $-3\mathbf{v}$.

We can consider linear combinations of vectors. Given any vectors $\mathbf{u}$ and $\mathbf{v}$, the linear combination

$$a\mathbf{u} + b\mathbf{v}$$

creates a new vector. Consider all vectors that can be created by picking different values of a and b. If $\mathbf{u}$ and $\mathbf{v}$ lie on a line, then these new vectors all lie on that same line. If $\mathbf{u}$ and $\mathbf{v}$ do not lie on a line, then the new vectors will fill ("spans") the entire two-dimensional space. In the first case, the vectors cannot fill the entire space. The two vectors can only span a one-dimensional space.

What is the length of a vector $\mathbf{u} = (u_1, u_2)$? By Pythagoras' Theorem, it is

$$\|\mathbf{u}\| = \sqrt{u_1^2 + u_2^2} = \sqrt{\mathbf{u} \cdot \mathbf{u}}.$$

The length of the vector is often called the "norm", and the relationship between the norm and $\mathbf{u} \cdot \mathbf{u}$ is

$$\|\mathbf{u}\|^2 = \mathbf{u} \cdot \mathbf{u}.$$

Exercise: let $\mathbf{u}=(4,2)$ and $\mathbf{v}=(-1,2)$. What are the lengths of the vectors? What are the lengths of the vectors $\mathbf{u}/\|\mathbf{u}\|$ and $\mathbf{v}/\|\mathbf{v}\|$.

The length of the vector $\mathbf{u}$ is also the distance between the origin and the point (u_1,u_2). The distance between any two points $\mathbf{u}$ and $\mathbf{v}$ is

$$\|\mathbf{u}-\mathbf{v}\|.$$

Suppose two vectors $\mathbf{u}=(u_1,u_2)$ and $\mathbf{v}=(v_1,v_2)$ are at right angles to each other. We say they are "orthogonal". We can show that their dot product must be zero. If $\mathbf{u}$ and $\mathbf{v}$ are at right angles to each other, then $\mathbf{u}-\mathbf{v}$ (or $\mathbf{v}-\mathbf{u}$) represents the hypotenuse. Then Pythagoras' Theorem says

$$\mathbf{u}\cdot\mathbf{u}+\mathbf{v}\cdot\mathbf{v}=(\mathbf{u}-\mathbf{v})\cdot(\mathbf{u}-\mathbf{v})$$

or

$$u_1^2+u_2^2+v_1^2+v_2^2=(u_1-v_1)^2+(u_2-v_2)^2.$$

Simplifying we get

$$u_1^2+u_2^2+v_1^2+v_2^2=u_1^2-2u_1v_1+v_1^2+u_2^2-2u_2v_2+v_2^2$$

$$u_1v_1+u_2v_2=0$$

$$\mathbf{u}\cdot\mathbf{v}=0.$$

We can generalize this result. We take it slow and first consider unit vectors (vectors with length one. Take the vectors $\mathbf{i}=(1,0)$ and $\mathbf{u}=(\cos\theta,\sin\theta)$. Both vectors are of unit length and the angle between them is θ. Also we have $\mathbf{i}\cdot\mathbf{u}=\cos\theta$. Now consider two unit length vectors $\mathbf{v}=(\cos\beta,\sin\beta)$ and $\mathbf{w}=(\cos\alpha,\sin\alpha)$ where $\beta-\alpha=\theta$. The angle between the two vectors is thus θ. We have that

$$\mathbf{u}.\mathbf{w}=\cos\beta\cos\alpha+\sin\beta\sin\alpha.$$

From the cosine formula in basic trigonometry, the right hand side is the same as $\cos(\beta-\alpha)$. Therefore

$$\mathbf{u}\cdot\mathbf{w}=\cos\theta.$$

What about non-unit vectors? If the vector has $\mathbf{u}$ has length $\|\mathbf{u}\|$, then the vector $\dfrac{\mathbf{u}}{\|\mathbf{u}\|}$ has length one. We therefore have

$$\frac{\mathbf{u}}{\|\mathbf{u}\|}\cdot\frac{\mathbf{w}}{\|\mathbf{w}\|}=\cos\theta.$$

or

$$\mathbf{u}\cdot\mathbf{w}=\|\mathbf{u}\|\|\mathbf{v}\|\cos\theta.$$

One of the consequences of this is that the vectors whose angles are between 0 and 90° have positive dot products, otherwise their dot products are negative. Another consequence is:

Theorem: Cauchy-Schwarz Inequality for two dimensional vectors

For any two-dimensional vectors $\mathbf{u}$ and $\mathbf{v}$, we have

$$|\mathbf{u} \cdot \mathbf{w}| \leq \|\mathbf{u}\| \|\mathbf{w}\|.$$

Proof: The result holds because $\cos\theta$ is always between -1 and 1.

This theorem (and generalizations of it) is one of the most important inequalities in all of mathematics. You will learn more about why this inequality is so important, though for many of us, we'll never get to fully appreciate its importance. For now, we note that the Cauchy-Schwarz inequality implies the following facts about distances in two-dimensions:

Theorem: Triangle Inequality for two-dimensional vectors

For any two-dimensional vectors $\mathbf{u}$ and $\mathbf{v}$, we have

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

(Can you see that this theorem says that the sum of two side of a triangle must be greater or equal than the length of the third side? Actually, equality holds only if the triangle is not actually a triangle, but a three points in a line.)

Proof: We have

$$\|\mathbf{u}+\mathbf{v}\|^2 = (\mathbf{u}+\mathbf{v}) \cdot (\mathbf{u}+\mathbf{v}) = \mathbf{u} \cdot \mathbf{u} + 2\mathbf{u} \cdot \mathbf{v} + \mathbf{v} \cdot \mathbf{v} = \|\mathbf{u}\|^2 + 2\mathbf{u} \cdot \mathbf{v} + \|\mathbf{v}\|^2.$$

Since $\mathbf{u} \cdot \mathbf{v}$ can be positive or negative, we have $\mathbf{u} \cdot \mathbf{v} \leq |\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\| \|\mathbf{v}\|$, which implies

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

Taking square roots gives us the result we are looking for.

The triangle inequality says that if you walk from point **x** to point **y** then to point **z** (in straight lines), the distance you travel will never be less than going from point **x** directly to point **z** (in a straight line), since

$$\|(\mathbf{x}-\mathbf{z})\| = \|(\mathbf{x}-\mathbf{y}) + (\mathbf{y}-\mathbf{z})\| \leq \|(\mathbf{x}-\mathbf{y})\| + \|(\mathbf{y}-\mathbf{z})\|.$$

<u>Some Geometry in Three Dimensions</u>

As with two-dimensional vectors, we can give three-dimensional vectors geometric meaning in three-dimensional space. You can imagine adding two three-dimensional vectors, scalar multiplication, and so on. What is the length of a three-dimensional vector $\mathbf{x} = (x_1, x_2, x_3)$? It takes two applications of Pythagoras' Theorem, but you can show that it is

$$\|\mathbf{u}\| = \sqrt{u_1^2 + u_2^2 + u_3^2} = \sqrt{\mathbf{u} \cdot \mathbf{u}}.$$

It is very convenient that we don't have to change either the norm or the dot product notation. Likewise, the vector $\mathbf{u} / \|\mathbf{u}\|$ has unit length.

What about the angle between two vectors? Suppose two vectors $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (v_1, v_2, v_3)$ are at right angles to each other. Again, we say they are "orthogonal", and again we can show that their dot product must be zero. If $\mathbf{u}$ and $\mathbf{v}$ are at right angles to each other, then $\mathbf{u} - \mathbf{v}$ (or $\mathbf{v} - \mathbf{u}$) represents the hypotenuse of a right-angled triangle. Then Pythagoras' Theorem says

$$\mathbf{u} \cdot \mathbf{u} + \mathbf{v} \cdot \mathbf{v} = (\mathbf{u} - \mathbf{v}) \cdot (\mathbf{u} - \mathbf{v})$$

or

$$u_1^2 + u_2^2 + u_3^2 + v_1^2 + v_2^2 + v_3^2 = (u_1 - v_1)^2 + (u_2 - v_2)^2 + (u_3 - v_3)^2 .$$

Simplifying we get

$$u_1^2 + u_2^2 + u_3^2 + v_1^2 + v_2^2 + v_3^2 = u_1^2 - 2u_1v_1 + v_1^2 + u_2^2 - 2u_2v_2 + v_2^2 + u_3^2 - 2u_3v_3 + v_3^2$$

$$u_1v_1 + u_2v_2 + u_3v_3 = 0$$

$$\mathbf{u} \cdot \mathbf{v} = 0 .$$

It is harder to visualize this, but it is also true that for two unit vectors in three-dimensional space, we have for unit vectors that

$$\mathbf{u} \cdot \mathbf{w} = \cos\theta$$

and for all vectors that

$$\frac{\mathbf{u}}{\|\mathbf{u}\|} \cdot \frac{\mathbf{w}}{\|\mathbf{w}\|} = \cos\theta .$$

or

$$\mathbf{u} \cdot \mathbf{w} = \|\mathbf{u}\| \|\mathbf{v}\| \cos\theta .$$

It follows immediately, for three-dimensional vectors, that:

Theorem: Cauchy-Schwarz Inequality for three-dimensional vectors

For any three-dimensional vectors $\mathbf{u}$ and $\mathbf{v}$, we have

$$|\mathbf{u} \cdot \mathbf{w}| \leq \|\mathbf{u}\| \|\mathbf{w}\| .$$

Proof: Because $\cos\theta$ is always between -1 and 1.

The triangle inequality thus continues to hold for three-dimensional vectors

Theorem: Triangle Inequality for three-dimensional vectors

For any three-dimensional vectors $\mathbf{u}$ and $\mathbf{v}$, we have

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

Proof: exactly the same as before.

As in two-dimensional space, the triangle inequality says that if you move (in three-dimensions) from point **x** to point **y** then to point **z** (in straight lines), the distance you travel will never be less than going from point **x** directly to point **z** (in a straight line), since again

$$\|(\mathbf{x} - \mathbf{z})\| = \|(\mathbf{x} - \mathbf{y}) + (\mathbf{y} - \mathbf{z})\| \leq \|(\mathbf{x} - \mathbf{y})\| + \|(\mathbf{y} - \mathbf{z})\| .$$

What about k-dimensional vectors ($k > 3$)?

Obviously we cannot give these geometric interpretations. However, suppose we define the norm of an *k*-dimensional vector $\mathbf{u} = (u_1, u_2, ..., u_k)$ as

$$\|\mathbf{u}\| = \sqrt{u_1^2 + u_2^2 + u_3^2 + ... + u_k^2} = \sqrt{\mathbf{u} \cdot \mathbf{u}}.$$

With this definition of a norm, it turns out that the Cauchy-Schwarz inequality continues to hold. In two-dimensions, we used geometric arguments to prove the inequality. We cannot do that here. However, there is a proof of the inequality that does not make use of any geometric ideas at all. This proof applies even to three- and two-dimensions (even one-dimensions!) and is completely independent of any geometric concepts.

Theorem: Cauchy-Schwarz Inequality for *k*-dimensional vectors

For any *k*-dimensional vectors $\mathbf{u}$ and $\mathbf{v}$, we have

$$|\mathbf{u} \cdot \mathbf{w}| \le \|\mathbf{u}\| \|\mathbf{w}\|.$$

Equality holds if and only if $\mathbf{u} = \alpha \mathbf{w}$ or one of the vectors is a zero vector.

Proof: If one of the vectors is a zero vector, then the result obviously holds trivially as an equality. If $\mathbf{u} = \alpha \mathbf{w}$, then $|\mathbf{u} \cdot \mathbf{w}| = |\alpha \mathbf{w} \cdot \mathbf{w}| = |\alpha| \|\mathbf{w}\|^2$ and also $\|\mathbf{u}\| \|\mathbf{w}\| = \alpha \|\mathbf{w}\|^2$, so again the result holds with equality. Suppose neither one is a zero vector, and that $\mathbf{u} \neq \alpha \mathbf{w}$ for any α. Then we have

$$0 < (\mathbf{u} - \alpha \mathbf{w}) \cdot (\mathbf{u} - \alpha \mathbf{w}) = \mathbf{u} \cdot \mathbf{u} - 2\alpha \mathbf{u} \cdot \mathbf{w} + \alpha^2 \mathbf{w} \cdot \mathbf{w}.$$

This holds for all α. Evaluating at one particular value, namely

$$\alpha = \frac{\mathbf{u} \cdot \mathbf{w}}{\mathbf{w} \cdot \mathbf{w}}$$

we get

$$0 < \mathbf{u} \cdot \mathbf{u} - 2\frac{\mathbf{u} \cdot \mathbf{w}}{\mathbf{w} \cdot \mathbf{w}} \mathbf{u} \cdot \mathbf{w} + \frac{(\mathbf{u} \cdot \mathbf{w})^2}{(\mathbf{w} \cdot \mathbf{w})^2} \mathbf{w} \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{u} - \frac{(\mathbf{u} \cdot \mathbf{w})^2}{\mathbf{w} \cdot \mathbf{w}}.$$

Rearranging the expression above gives

$$(\mathbf{u} \cdot \mathbf{w})^2 \le (\mathbf{u} \cdot \mathbf{u})(\mathbf{w} \cdot \mathbf{w}).$$

Taking square roots gives the desired result.

The fact that the Cauchy-Schwarz inequality holds in *k*-dimensions means that the triangle inequality also holding for *k*-dimension space, since the triangle inequality is an immediate implication of the Cauchy-Schwarz inequality.

Theorem: Triangle Inequality for *k*-dimensional vectors

For any *k*-dimensional vectors $\mathbf{u}$ and $\mathbf{v}$, we have

$$\|\mathbf{u} + \mathbf{v}\| \le \|\mathbf{u}\| + \|\mathbf{v}\|.$$

Proof: exactly the same as before.

This leads to several interesting observations. Note that we called $\|\mathbf{x}\|$ a 'norm', not a length, since in *k*-dimensional space, it is not clear what "length" means. However, the norm behaves in a way that is very similar to that of length and distances.

Consider the distance between two points **u** and **v** (either in two- or three-dimensional space). What are the essential properties of this 'geometric distance'? First it is always positive

$$\|\mathbf{u}-\mathbf{v}\| \geq 0.$$

Second, it is symmetric – moving from **u** to **v**, or from **v** to **u**, it's the same distance, i.e.

$$\|\mathbf{u}-\mathbf{v}\| = \|\mathbf{v}-\mathbf{u}\|.$$

Third, the triangle inequality must hold. If I go from **u** to **w**, then **w** to **v**, my distance travelled will not be less than if I travelled from **u** to **v** directly:

$$\|\mathbf{u}-\mathbf{v}\| \leq \|\mathbf{u}-\mathbf{w}\| + \|\mathbf{w}-\mathbf{v}\|.$$

However, we have just proven that these properties also hold for norms of *k*-dimensional vectors. So even though we cannot interpret $\|\mathbf{u}\|$ as a geometric length, or $\|\mathbf{u}-\mathbf{v}\|$ as a geometric distance for *k*-dimensional vectors, they do in fact behave very much like lengths and distances and can be treated as such, and we can use 2-d and 3-d geometric intuitions to help us think about abstract *k*-dimensional spaces.

Furthermore, since the Cauchy-Schwarz inequality says that

$$\frac{|\mathbf{u}\cdot\mathbf{w}|}{\|\mathbf{u}\|\|\mathbf{w}\|} \leq 1.$$

Since in 2-d and 3-d spaces we have

$$\frac{\mathbf{u}}{\|\mathbf{u}\|}\cdot\frac{\mathbf{w}}{\|\mathbf{w}\|} = \cos\theta,$$

we shall <u>define</u> the angle between two *k*-dimensional vectors as

$$\theta = \cos^{-1}\left(\frac{\mathbf{u}}{\|\mathbf{u}\|}\cdot\frac{\mathbf{w}}{\|\mathbf{w}\|}\right)?$$

This definition (even though we do not have a geometric notion of an angle in n-dimensional space) allows us to continue to say that

$$\frac{\mathbf{u}}{\|\mathbf{u}\|}\cdot\frac{\mathbf{w}}{\|\mathbf{w}\|} = \cos\theta,$$

The Cauchy-Schwarz guarantees also guarantees that we still have $-1 \leq \cos\theta \leq 1$, and it allows us to say that two *k*-dimension vectors are orthogonal (with all geometric intuitions intact) if $\mathbf{u}\cdot\mathbf{w} = 0$.

Exercises

1. Write the Cauchy-Schwarz inequality

$$|\mathbf{u}\cdot\mathbf{w}|\leq\|\mathbf{u}\|\|\mathbf{w}\|$$

using summation notation, where $\mathbf{u}$ and $\mathbf{v}$ are n-dimensional vectors.

2. Given observations $x_1, x_2, ..., x_n$ and $y_1, y_2, ..., y_n$ on two variables, the sample covariance is

$$\text{sample cov}[x_i, y_i] = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y}).$$

Let $\tilde{\mathbf{x}} = (x_1-\overline{x}, x_2-\overline{x}, ..., x_n-\overline{x})$ and $\tilde{\mathbf{y}} = (y_1-\overline{y}, y_1-\overline{y}, ..., y_1-\overline{y})$. Write down the sample covariance in terms of the dot product of $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$. Thus, the sample covariance is zero if the two vectors $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$ are orthogonal. How would you write the sample variance in terms of the dot product?

The sample correlation is defined as

$$\text{sample corr}[x_i, y_i] = \frac{\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\overline{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\overline{y})^2}}$$

Explain why the sample correlation is always between -1 and 1.