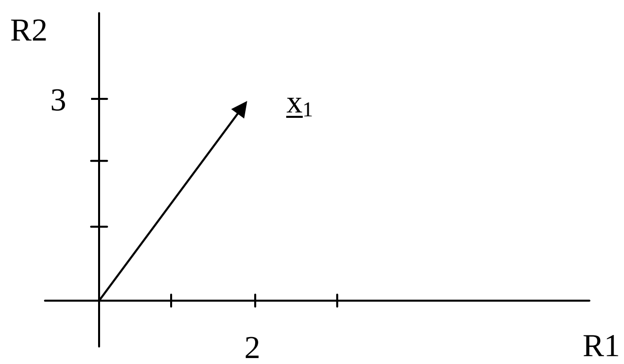
The Geometry of Vectors

A vector can be represented geometrically by using a coordinate system where each element in the vector corresponds to a distance along one of the reference axes defining the system.

Ex/ $\underline{x}_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, where 2 is the distance on the first reference axis (R1) and 3 is the distance on the second reference axis (R2).



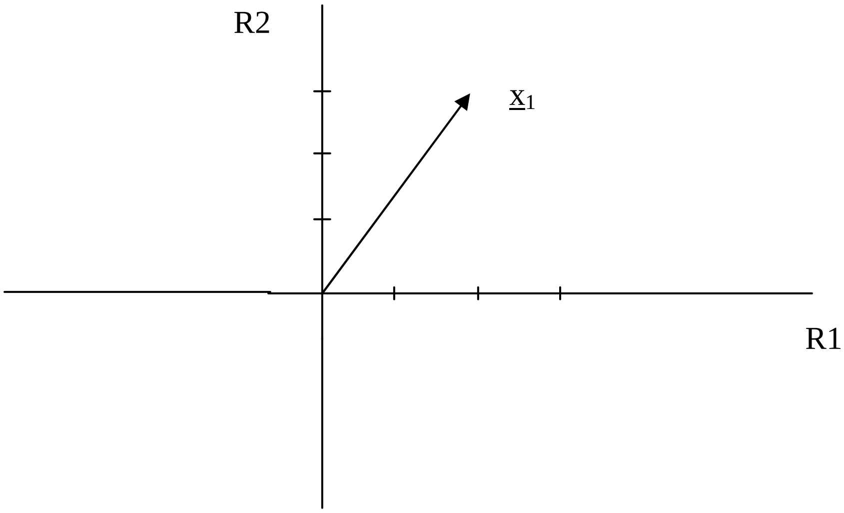
The arrow at the end of the vector suggests the direction of the vector – it is the end (terminus) of the vector.

Note:

1. this is a two-dimensional space
2. the reference axes are at right angles, representing an orthogonal reference system
3. the reference axes intersect at the origin of the space

Consider three additional vectors

$\underline{x}_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, $\underline{x}_3 = \begin{bmatrix} -3 \\ 2 \end{bmatrix}$, $\underline{x}_4 = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$ Add these to the coordinate system.



Note that these four vectors can be described in terms of (1) length and (2) direction. Differences between vectors can also be described in terms of the angles between them.

In the general form, a *n*-element vector exists in a *n*-dimensional space.

Vector Length

The length of a vector can be obtained through the use of the Pythagorean theorem. This theorem applies because the reference axes are at right angles. In this system, each vector can be considered the hypotenuse of a right triangle. Because of this, the sum of the squares of the coordinates is equal to the square of the length of the vector.

$a^2 + b^2 = c^2$

Length of $\underline{x}_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ is the square root of $2^2 + 3^2 = \sqrt{13} = 3.61$.

$L(\underline{x}) = \sqrt{\underline{x}' \ \underline{x}}$ so that $L(\underline{x}_2) = 2.24$, $L(\underline{x}_3) = 3.61$, & $L(\underline{x}_4) = 5.00$

Vector Angles

The cosine of an angle θ is useful because it is defined as the length of the side of the triangle which is adjacent to θ divided by the length of the hypotenuse of the triangle.

$$\cos\theta = \frac{L(A)}{L(H)}$$

The cosine of an angle is zero when two vectors are separated by a 90° angle or are orthogonal. This indicates that one vector contains no information about another vector.

The cosine of an angle is one when two vectors are collinear, indicating that the information in the two vectors is redundant.

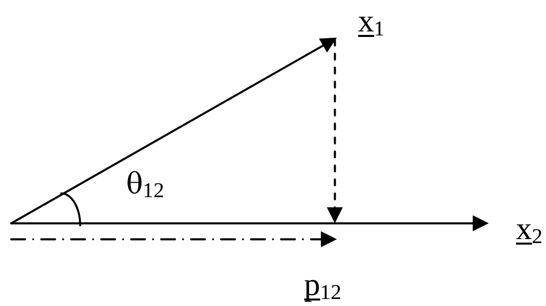
There is a relation between the scalar product of two vectors and the angle between them.

$$\underline{x}'_1\underline{x}_2 = L(\underline{x}_1)\,L(\underline{x}_2)\cos\theta_{12}$$

$$\cos\theta_{12} = \frac{\underline{x}'_1\underline{x}_2}{L(\underline{x}_1)L(\underline{x}_2)}$$

Vector Orthogonal Projection

Orthogonal projection of $\underline{x}_1$ on $\underline{x}_2$ occurs by dropping a perpendicular line from the terminus of $\underline{x}_1$ to intersect with $\underline{x}_2$.



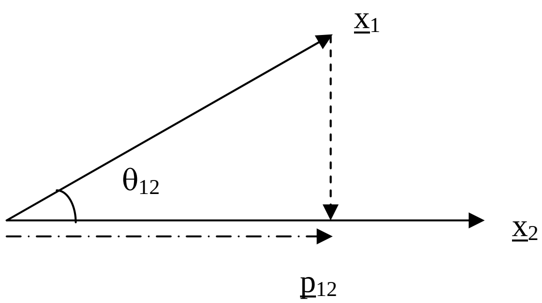
The result is a vector $\underline{p}_{12}$ which is collinear with $\underline{x}_2$ but has a different length.

Relying on the trig function, $\text{Cos}\,\theta = \frac{A}{H} = \frac{L(\underline{p}_{12})}{L(\underline{x}_1)}$.

The length of $\underline{p}_{12}$ is defined by $L(\underline{p}_{12}) = L(\underline{x}_1)\,\text{Cos}\,\theta_{12}$.

Computing the Length of the Projection

Given the relation: Cos $\theta_{12} = \dfrac{\underline{x}'_1 \underline{x}_2}{L(\underline{x}_1)L(\underline{x}_2)}$ and Cos $\theta = \dfrac{A}{H} = \dfrac{L(\underline{p}_{12})}{L(\underline{x}_1)}$.



Next, solve for the length of the projection: $L(\underline{p}_{12}) = L(\underline{x}_1)$ Cos θ_{12}.

$\underline{x}_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, $\underline{x}_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, $\underline{x}_3 = \begin{bmatrix} -3 \\ 2 \end{bmatrix}$, $\underline{x}_4 = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$

In our previous example of vectors compute the following:

$L(\underline{p}_{12})$. First solve for Cos $\theta_{12} = \dfrac{2 \cdot 2 + 3 \cdot -1}{\sqrt{2^2 + 3^2}\sqrt{2^2 + -1^2}} = \dfrac{1}{\sqrt{13 \cdot 5}} = .124$ ~ 83° Angle

$L(\underline{p}_{12}) = L(\underline{x}_1)$ Cos $\theta_{12} = \sqrt{13} \cdot (.124) = .447$

$L(\underline{p}_{14})$. First solve for Cos $\theta_{14} = \dfrac{2 \cdot 4 + 3 \cdot 3}{\sqrt{2^2 + 3^2}\sqrt{4^2 + 3^2}} = \dfrac{17}{\sqrt{13 \cdot 25}} = .943$ ~ 19° Angle

$L(\underline{p}_{14}) = L(\underline{x}_1)$ Cos $\theta_{14} = \sqrt{13} \cdot (.943) = 3.4$

$L(\underline{p}_{42})$. First solve for Cos θ_{42} =

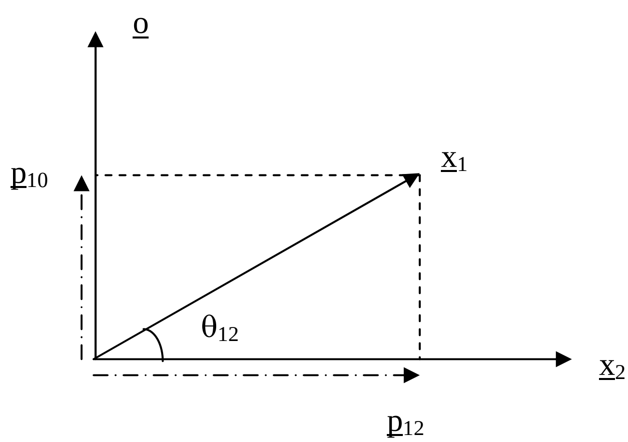
$L(\underline{p}_{42}) = L(\underline{x}_4)$ Cos θ_{42} =

$L(\underline{p}_{23})$. First solve for Cos θ_{23} =

$L(\underline{p}_{23}) = L(\underline{x}_2)$ Cos θ_{23} =

Vector Orthogonal Decomposition

Orthogonal projection of a vector results in a geometric decomposition of the vector into two additive components.



The first projection is the same orthogonal projection from the earlier example, $\underline{p}_{12}$. The second projection, $\underline{p}_{1O}$, is the projection of $\underline{x}_1$ onto vector $\underline{o}$ that is by definition orthogonal to $\underline{x}_2$.

The lengths of the two projections correspond to the lengths of the sides of the right triangle where the hypotenuse is $\underline{x}_1$.

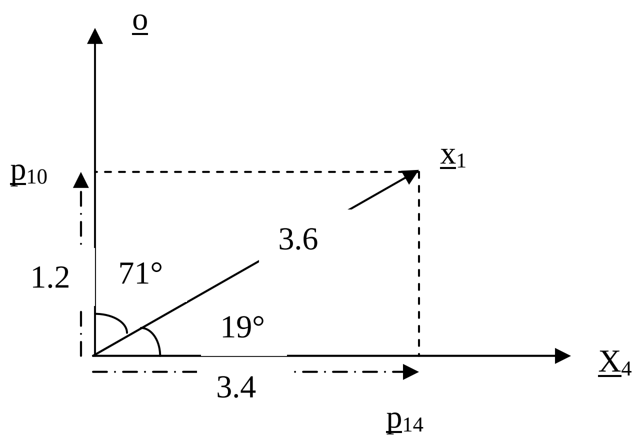
$L(\underline{p}_{1O}) = L(\underline{x}_1) \text{ Cos } \theta_{1O}$

Example: $\underline{x}_1$ and $\underline{x}_4$

Since $\theta_{14} = 19°$, $\theta_{1O} = 90 – 19 = 71°$. So $\text{Cos } \theta_{1O} = \text{Cos } 71° = 0.326$

$L(\underline{p}_{1O}) = (3.61)(0.326) = 1.18$

The final system is fully described here:



We can use the Pythagorean theorem to verify the computation of the two projections. Since the square of the hypotenuse of a right triangle is the sum of the squares of the sides, we can check our work:

$$[L(\underline{x}_1)]^2 = [L(\underline{p}_{12})]^2 + [L(\underline{p}_{1O})]^2$$

$$3.61^2 = (3.40)^2 + (1.18)^2 = 11.6 + 1.4 = 13.0$$

Deviation Scores and Standard Deviations

In statistics, we employ deviation scores often and some even use deviation scores regularly rather than raw scores.

Recognize that the length of a vector is the square root scalar product (inner product) of the vector. If we were to compute the length of a deviation score vector, the length would be the square root of the sum of squared deviations; when divided by $\sqrt{n}$, this would provide the standard deviation. So there is a direct relation between the length of deviation vectors and the standard deviation.

$$L(\underline{d}) = \sqrt{n}\, s \quad \text{or alternatively,} \quad s = \frac{1}{\sqrt{n}} L(\underline{d}) .$$

Ex/ Consider data on GPA (X_1) and SAT (X_2) scores for 10 students.

$$\underline{x}_1 = \begin{bmatrix} 3.8 \\ 2.4 \\ 2.6 \\ 3.1 \\ 1.9 \\ 1.7 \\ 2.5 \\ 2.4 \\ 3.5 \\ 3.1 \end{bmatrix} \quad \underline{x}_2 = \begin{bmatrix} 760 \\ 710 \\ 680 \\ 730 \\ 420 \\ 410 \\ 620 \\ 630 \\ 720 \\ 670 \end{bmatrix}$$

We can compute means for both: $\bar{X}_1 = 2.7$, $\bar{X}_2 = 635$.

By subtracting a vector of means from each vector of data, we can create deviation vectors.

Recall that $\underline{d} = \left(X - \bar{X}\right)$ and $L(\underline{d}) = \sqrt{\sum\left(X - \bar{X}\right)^2} = \sqrt{\underline{d}'\underline{d}}$

$$\underline{d}_1 = \begin{bmatrix} 1.1 \\ -.3 \\ -.1 \\ .4 \\ -.8 \\ -1 \\ -.2 \\ -.3 \\ .8 \\ .4 \end{bmatrix} \qquad \underline{d}_2 = \begin{bmatrix} 125 \\ 75 \\ 45 \\ 95 \\ -215 \\ -225 \\ -15 \\ -5 \\ 85 \\ 35 \end{bmatrix} \qquad L(\underline{d}_1) = \sqrt{4.04} = 2.01 \quad L(\underline{d}_2) = \sqrt{137850} = 371$$

$$s_1 = \frac{1}{\sqrt{10}}\ 2.01 = 0.64 \qquad s_2 = \frac{1}{\sqrt{10}}\ 371 = 117$$

Vector Correlation and Separation

Because the correlation can be expressed in terms of deviation scores, we can see that

$$r = \frac{Cov_{xy}}{s_x s_y} = \frac{\frac{\sum xy}{n}}{\sqrt{\frac{\sum x^2}{n}}\sqrt{\frac{\sum y^2}{n}}} = \frac{\frac{d'_x d_y}{n}}{\sqrt{\frac{d'_x d_x}{n}}\sqrt{\frac{d'_y d_y}{n}}} = \frac{\underline{d}'_x \underline{d}_y}{\sqrt{\underline{d}'_x \underline{d}_x}\sqrt{\underline{d}'_y \underline{d}_y}} = \frac{\underline{d}'_x \underline{d}_y}{L(\underline{d}_x)L(\underline{d}_y)}$$

Recall that $\cos\theta_{12} = \frac{\underline{x}'_1 \underline{x}_2}{L(\underline{x}_1)L(\underline{x}_2)}$, so that when we are engaging deviation vectors, $\cos\theta_{12} = r_{12}$.

Perfect positive correlation ($r_{12} = 1$) indicates that deviation vectors are collinear ($\theta_{12} = 0°$).

Perfect negative correlation ($r_{12} = -1$) indicates that deviation vectors are opposite ($\theta_{12} = 180°$).

Perfect lack of correlation ($r_{12} = 0$) indicates that the deviation vectors are orthogonal ($\theta_{12} = 90°$).

We can compute the correlation between our scores on GPA and SAT:

$\underline{x_1}$	$\underline{x_2}$	$\underline{d_1}$	$\underline{d_1}^2$	$\underline{d_2}$	$\underline{d_2}^2$	$\underline{d_1 d_2}$
3.8	760	1.1	1.21	125	15625	137.5
2.4	710	-0.3	0.09	75	5625	-22.5
2.6	680	-0.1	0.01	45	2025	-4.5
3.1	730	0.4	0.16	95	9025	38
1.9	420	-0.8	0.64	-215	46225	172
1.7	410	-1.0	1.00	-225	50625	225
2.5	620	-0.2	0.04	-15	225	3
2.4	630	-0.3	0.09	-5	25	1.5
3.5	720	0.8	0.64	85	7225	68
3.1	670	0.4	0.16	35	1225	14
mean=2.7	mean=635		sum =4.04		sum=137850	632

$$r_{12} = \frac{632}{(2.01)(371)} = 0.8475$$

This correlation corresponds to an angle of approximately 32° of separation in deviation score vectors for GPA and SAT in the 10-dimensional space from our 10 students.

Vector Orthogonal Decomposition and Bivariate Regression

We have already reviewed the typical regression model where variation in one variable (Y) is explained by another variable (X), employing a linear model:

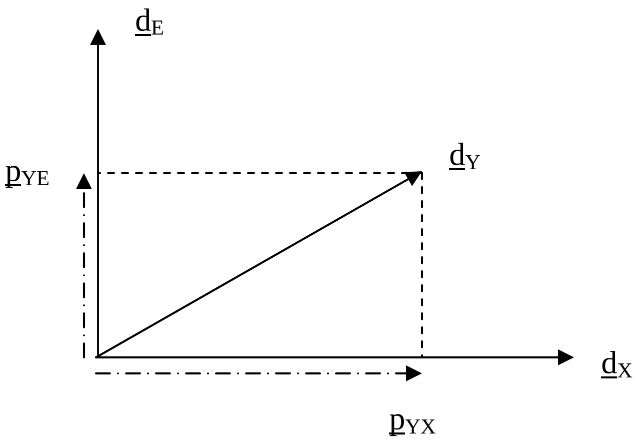
$$Y_i = b_0 + b_1 X_i + e_i$$

We also saw how this model partitions the total sums of squares of Y into two additive components, sums of squares regression and sums of squares error:

$$SS_T = SS_R + SS_E$$

This partitioning of sums of squares can be represented geometrically using deviation score vectors $\underline{d}_y$ and $\underline{d}_x$. Regression of Y on X is analogous to the orthogonal decomposition of $\underline{d}_y$.

To do this, we must complete the orthogonal projection of vector $\underline{d}_y$ on vector $\underline{d}_x$ and a second vector that is orthogonal to $\underline{d}_x$ which we call $\underline{d}_E$. (similar to $\underline{o}$ used earlier).



The first projection involves projecting $\underline{d}_Y$ on $\underline{d}_X$ resulting in $\underline{p}_{YX}$, which is referred to as the component of $\underline{d}_Y$ along $\underline{d}_X$. The square of the length of $\underline{p}_{YX}$ is analogous to the sums of squares regression.

$$[L(\underline{p}_{YX})]^2 = SS_R.$$

The second projection consist of the component of $\underline{d}_Y$ that is uncorrelated with $\underline{d}_X$, which is the projection of $\underline{d}_Y$ on the error vector, $\underline{d}_E$. The square of the length of $\underline{p}_{YE}$ is analogous to the sums of squares error; the length represents the component of $\underline{d}_Y$ that remains after the component related to $\underline{d}_X$ has been removed.

$$[L(\underline{p}_{YE})]^2 = SS_E.$$

We can then apply the Pythagorean theorem to this system to obtain:

$$[L(\underline{d}_Y)]^2 = [L(\underline{p}_{YX})]^2 + [L(\underline{p}_{YE})]^2 \rightarrow SS_T = SS_R + SS_E$$

We could use our GPA and SAT data to provide an example of these computations. Suppose we want to predict SAT ($\underline{d}_Y$) from GPA ($\underline{d}_X$), employing deviation scores.

The lengths of GPA ($\underline{d}_X$) and SAT ($\underline{d}_Y$) reflect differences in standard deviations.

$L(\underline{d}_X) = \sqrt{4.04} = 2.01 \quad L(\underline{d}_Y) = \sqrt{137850} = 371.3$

The cosine of the angle between the two vectors represents the correlation (0.8475).

The projections would have lengths computed by:

$L(\underline{p}_{YX}) = L(\underline{d}_Y) \text{ Cos } \theta_{YX} = (371.3)(0.8475) = 314.68$

$L(\underline{p}_{YE}) = L(\underline{d}_Y) \text{ Cos } \theta_{YE} = (371.3)(0.5299) = 196.75$

note: since the angle of separate between Y and X is 32^0, the angle between Y and E must be approximately 58^0, corresponding to a cosine value of 0.5299.

The partitioning of the sums of squares is then:

$(371.3)^2 = (314.68)^2 + (196.75)^2$

$137864 = 99024 + 38711$ *approximately equal within rounding error*

And, of course, from these sums of squares we can estimate R^2 and R.

$R^2 = 99024/137864 = 0.718 \rightarrow R = 0.8475$