Problem 1: Geometry with dot products

(a) Using that perpendicularity is governed by the dot products being equal to 0, find a nonzero vector in $\mathbf{R}^3$ that is perpendicular to $\mathbf{v} = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$. Then find another that is not a scalar multiple of that one.

(b) Find an equation in x, y, z that characterizes when $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ is perpendicular to $\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$. What does this collection of vectors look like?

(c) (**Extra**) What does the collection of nonzero vectors $\mathbf{w} = \begin{bmatrix} x \\ y \end{bmatrix}$ making an angle of at most $60°$ against $\mathbf{v} = \begin{bmatrix} 3 \\ -4 \end{bmatrix}$ look like? Using the relation of dot products and cosines to describe this region with a pair of conditions of the form $ax^2 + bxy + cy^2 \geq 0$ and $y \leq (3/4)x$ (away from the origin).

Solution: The perpendicularity against $\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$ is the condition $0 = 2x - y + z$. This answers the first part of (b), and the collection of all such vectors is a plane through the origin (we visualize the directions perpendicular to the line through $\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$). For (a), we can pick whatever x and y we like (not both 0) and then set z to be $y - 2x$ to enforce perpendicularity. So two such vectors are $\begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$.

For (c), it looks like a sector: the region between two half-lines emanating from the origin (at a total angle of $120°$). Since $\cos 60° = 1/2$, the condition (for nonzero $\mathbf{w}$) is

$$\frac{1}{2} \leq \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\|\|\mathbf{w}\|} = \frac{3x - 4y}{5\sqrt{x^2 + y^2}},$$

or equivalently

$$\sqrt{x^2 + y^2} \leq (2/5)(3x - 4y).$$

This forces $3x - 4y \geq 0$, or in other words $y \leq (3/4)x$, in which case it is harmless to square both sides to arrive at the equivalent inequality $x^2 + y^2 \leq (4/25)(9x^2 - 24xy + 16y^2)$, or in other words $25x^2 + 25y^2 \leq 36x^2 - 96xy + 64y^2$, which is to say $11x^2 - 96xy + 39y^2 \geq 0$ (away from the origin since we needed $\mathbf{w} \neq \mathbf{0}$ to make sense of the angle, though allowing the origin for this final inequality simply inserts the vertex of the sector).

Problem 2: Algebra with dot products

(a) For $\mathbf{a} = \begin{bmatrix} 4 \\ -2 \\ 3 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 1 \\ 5 \\ -2 \end{bmatrix}$, and $\mathbf{c} = \begin{bmatrix} 6 \\ -4 \\ -1 \end{bmatrix}$ show that $\mathbf{a} \cdot (\mathbf{b} - \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} - \mathbf{a} \cdot \mathbf{c}$.

(b) Give an example of 2-vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ for which $(\mathbf{a} \cdot \mathbf{b})\mathbf{c} \neq (\mathbf{a} \cdot \mathbf{c})\mathbf{b}$. (Hint: what if $\mathbf{b}$ and $\mathbf{c}$ are not on the same line through the origin?)

(c) (**Extra**) Explain in terms of variables why $\mathbf{v} \cdot (\mathbf{w}_1 + \mathbf{w}_2) = \mathbf{v} \cdot \mathbf{w}_1 + \mathbf{v} \cdot \mathbf{w}_2$ for any 3-vectors $\mathbf{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, $\mathbf{w}_1 = \begin{bmatrix} a_1 \\ b_1 \\ c_1 \end{bmatrix}$, and $\mathbf{w}_2 = \begin{bmatrix} a_2 \\ b_2 \\ c_2 \end{bmatrix}$. If you then replace $\mathbf{v}$ with $\mathbf{v}_1 + \mathbf{v}_2$ for 3-vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ and apply another instance of the same general identity, why does it follow without any extra work with algebra in vector entries that

$$(\mathbf{v}_1 + \mathbf{v}_2) \cdot (\mathbf{w}_1 + \mathbf{w}_2) = \mathbf{v}_1 \cdot \mathbf{w}_1 + \mathbf{v}_1 \cdot \mathbf{w}_2 + \mathbf{v}_2 \cdot \mathbf{w}_1 + \mathbf{v}_2 \cdot \mathbf{w}_2?$$

(This is showing the analogue for vectors of the fact for numbers that the distributive law $r(s+t) = rs+rt$ is what makes the identity $(a+b)(c+d) = ac+ad+bc+bd$ hold, since $(a+b)(c+d) = (a+b)c+(a+b)d = ac+bc+ad+bd$.)

Do your arguments work for n-vectors for any n?

(d) For n-vectors $\mathbf{w}_1$ and $\mathbf{w}_2$, verify that $\|\mathbf{w}_1+\mathbf{w}_2\|^2 = \|\mathbf{w}_1\|^2+2(\mathbf{w}_1\cdot\mathbf{w}_2)+\|\mathbf{w}_2\|^2$ by using the relation $\|\mathbf{w}\|^2 = \mathbf{w}\cdot\mathbf{w}$ and general properties of dot products as stated in (c) (even if you didn't do (c)), *not* by writing out big formulas for lengths and dot products in terms of vector entries.

Solution:

(a) We compute

$$\mathbf{a} \cdot (\mathbf{b} - \mathbf{c}) = \begin{bmatrix} 4 \\ -2 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} -5 \\ 9 \\ -1 \end{bmatrix} = -20 - 18 - 3 = -41$$

and

$$\mathbf{a} \cdot \mathbf{b} - \mathbf{a} \cdot \mathbf{c} = \begin{bmatrix} 4 \\ -2 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 5 \\ -2 \end{bmatrix} - \begin{bmatrix} 4 \\ -2 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 6 \\ -4 \\ -1 \end{bmatrix} = (4 - 10 - 6) - (24 + 8 - 3) = -12 - 29 = -41.$$

(b) There are many options. Take $\mathbf{b} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{c} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, so the two sides point along different lines as long as they're nonzero. So take $\mathbf{a}$ not perpendicular to either one, say $\mathbf{a} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Then the left side is $\mathbf{c}$ and the right side is $\mathbf{b}$.

(c) Writing $\mathbf{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, $\mathbf{w}_i = \begin{bmatrix} a_i \\ b_i \\ c_i \end{bmatrix}$ we have

$$\mathbf{v} \cdot (\mathbf{w}_1 + \mathbf{w}_2) = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \cdot \begin{bmatrix} a_1 + a_2 \\ b_1 + b_2 \\ c_1 + c_2 \end{bmatrix} = x(a_1 + a_2) + y(b_1 + b_2) + z(c_1 + c_2)$$

and

$$\mathbf{v}\cdot\mathbf{w}_1+\mathbf{v}\cdot\mathbf{w}_2 = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ b_1 \\ c_1 \end{bmatrix} + \begin{bmatrix} x \\ y \\ z \end{bmatrix} \cdot \begin{bmatrix} a_2 \\ b_2 \\ c_2 \end{bmatrix} = (xa_1+yb_1+zc_1)+(xa_2+yb_2+zc_2) = xa_1+xa_2+yb_1+yb_2+zc_1+zc_2,$$

and these two outcomes are equal because of the distributive law (i.e., $r(s+t) = rs+rt$). For this reason, the same calculation works for n-vectors for every n (each vector entry is treated on its own – no interaction among entries in different positions – so it doesn't matter how many of them there are).

In the case $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$, we then have

$$(\mathbf{v}_1 + \mathbf{v}_2) \cdot (\mathbf{w}_1 + \mathbf{w}_2) = (\mathbf{v}_1 + \mathbf{v}_2) \cdot \mathbf{w}_1 + (\mathbf{v}_1 + \mathbf{v}_2) \cdot \mathbf{w}_2 = (\mathbf{v}_1 \cdot \mathbf{w}_1 + \mathbf{v}_2 \cdot \mathbf{w}_1) + (\mathbf{v}_1 \cdot \mathbf{w}_2 + \mathbf{v}_2 \cdot \mathbf{w}_2),$$

and the right side is the same as the desired expression up to rearranging terms in this sum of 4 numbers.

(d) We use the distributive law for dot product over vector addition (as discussed in (b)):

$$\|\mathbf{w}_1 + \mathbf{w}_2\|^2 = (\mathbf{w}_1 + \mathbf{w}_2) \cdot (\mathbf{w}_1 + \mathbf{w}_2) = \mathbf{w}_1 \cdot \mathbf{w}_1 + \mathbf{w}_1 \cdot \mathbf{w}_2 + \mathbf{w}_2 \cdot \mathbf{w}_1 + \mathbf{w}_2 \cdot \mathbf{w}_2,$$

and the right side is what we want since the outer terms are length squared and the two middle terms agree and hence add up to $2(\mathbf{w}_1 \cdot \mathbf{w}_2)$.

Problem 3: A correlation coefficient

Consider the collection of 5 data points: $(-2, 5)$, $(-1, 3)$, $(0, 0)$, $(1, -2)$, $(2, -6)$.

(a) Plot the points to see if they look close to a line.

(b) Compute the compute correlation coefficient exactly. Plug that into a calculator to approximate it to three decimal digits to see if its nearness to ± 1 fits well with the visual quality of fit of the line to the data plot in (a).

Solution:

(a) A plot of the data, as in Figure 1, shows it is reasonably close to a line with negative slope.

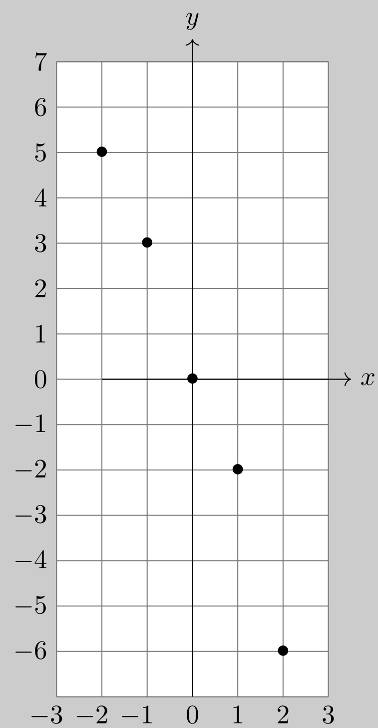

Figure 1: The data appears to fit a line of negative slope quite well.

(b) The initial data vectors are $\mathbf{X} = \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{bmatrix}$ and $\mathbf{Y} = \begin{bmatrix} 5 \\ 3 \\ 0 \\ -2 \\ -6 \end{bmatrix}$.

The correlation coefficient is

$$r = \frac{\mathbf{X} \cdot \mathbf{Y}}{\|\mathbf{X}\| \|\mathbf{Y}\|},$$

and we have

$$\mathbf{X} \cdot \mathbf{Y} = -27, \quad \|\mathbf{X}\| = \sqrt{10}, \quad \|\mathbf{Y}\| = \sqrt{74},$$

so $r = -27/(\sqrt{10}\sqrt{74}) \approx -0.992$. This is extremely close to -1, as we would expect since the data is seen by inspection to look very close to a line of negative slope (though the actual negative slope is not -1, as the line is quite steep).

Problem 4: More convex combinations (Extra)

(a) For the 2-vectors $\mathbf{a} = \begin{bmatrix}1\\0\end{bmatrix}, \mathbf{b} = \begin{bmatrix}0\\1\end{bmatrix}, \mathbf{c} = \begin{bmatrix}0\\0\end{bmatrix}$, describe the set of all possible vectors $r\mathbf{a} + s\mathbf{b} + t\mathbf{c}$ where $r + s + t = 1$ with $0 \leq r, s, t \leq 1$. Which points in your description correspond to the case $t = 0$. How about $s = 0$? Or $r = 0$? (Hint: plot points for a variety of triples $(r, s, t) = (r, s, 1 - r - s)$ with $0 \leq r, s, 1 - (r + s) \leq 1$.)

(b) Try the same using the 3-vectors $\mathbf{a} = \begin{bmatrix}1\\0\\0\end{bmatrix}, \mathbf{b} = \begin{bmatrix}0\\1\\0\end{bmatrix}$, and $\mathbf{c} = \begin{bmatrix}0\\0\\1\end{bmatrix}$. (Hint: first sketch the points you get with $t = 0$, then with $s = 0$, then with $r = 0$, and finally with $r = s = t = 1/3$.)

(c) Can you explain why your description in (a) applies to any three 2-vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ not on a common line? Use whatever physical or mathematical idea comes to mind. (Here is one approach: for $0 \leq t < 1$ check the equality $r\mathbf{a} + s\mathbf{b} + t\mathbf{c} = (1 - t)\mathbf{d}_{r,s} + t\mathbf{c}$ with a convex combination on the right where $\mathbf{d}_{r,s}$ is defined to be the convex combination $(r/(r + s))\mathbf{a} + (s/(r + s))\mathbf{b}$; this algebra works because $r + s = 1 - t > 0$. Interpret these convex combinations geometrically.)

(d) Is there a version for a triple of 3-vectors not all on a common line in space? Can you explain why it works?

Solution:

(a) This gives all points in the triangle bounded by the coordinate axes and the line $x + y = 1$; setting one of r, s, t to be 0 gives the side of the triangle opposite the corner corresponding to the parameter set to be 0.

(b) Setting one of the parameters to be 0 gives the respective segments $x + y = 1$ in the xy-plane ($z = 0$), $x + z = 1$ in the xz-plane ($y = 0$), and $y + z = 1$ in the yz-plane ($x = 0$). Varying the parameters more generally gives the points in the triangle with those segments as edges (and the points $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$ as corners).

(c) We always get the triangle with corners at $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$; when all of r, s, t are positive then we're on the interior and if any of them equal 0 then we're on an edge (and if two of them are 0 – so the third parameter is equal to 1 – then we're at a corner).

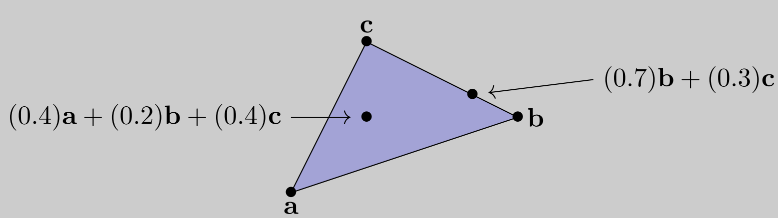

Figure 2: A triangle with vertices $\mathbf{a}, \mathbf{b}, \mathbf{c}$. We have shown two convex combinations of the vertices.

To explain this, given $0 \leq t < 1$ we note that the values $r, s \geq 0$ with $r+s = 1-t$ make the ratios $r/(r+s) = r/(1-t)$ and $s/(r+s) = s/(1-t) = 1-r/(r+s)$ go through exactly the pairs $(q, 1-q)$ with $0 \leq q \leq 1$. The points $\mathbf{d}_{r,s}$ for such varying r and s therefore account for exactly the points on the edge E joining $\mathbf{a}$ to $\mathbf{b}$ without repetition. Thus, $(1-t)\mathbf{d}_{r,s} + t\mathbf{c}$ is the point on the segment joining $\mathbf{d}_{r,s}$ to the other vertex $\mathbf{c}$ whose distance along the segment from $\mathbf{d}_{r,s}$ at a proportion t of the entire length of that segment (think about $t = 0$). So we are sweeping out the triangle using all of the line segments joining one vertex $\mathbf{c}$ to each of the points on the opposite edge E, with t keeping proportional track of where along such a segment a point is located distance-wise from the endpoint on E.

(d) The argument in the solution to (c) works without change for 3-vectors, since all of the reasoning in terms of the geometry of a triangle (now in space, rather than in a plane) continues to hold, and likewise for the geometric meaning of a convex combination $t\mathbf{v} + (1-t)\mathbf{w}$ for $\mathbf{v}, \mathbf{w}$ in $\mathbf{R}^3$ rather than in $\mathbf{R}^2$ (much as the interpretation of such convex combinations in terms of being at a point some proportion of the distance along a line segment works for 3-vectors as well as it does for 2-vectors).

