Peeter Joot
peeterjoot@pm.me

Revisiting linear solution systems using the wedge product

1.1 Exact system.

Recall that we can use the wedge product to solve linear systems. For example, assuming that $\mathbf{a}, \mathbf{b}$ are not colinear, the system

$$x\mathbf{a} + y\mathbf{b} = \mathbf{c}, \tag{1.1}$$

if it has a solution, can be solved for x and y by wedging with $\mathbf{b}$, and $\mathbf{a}$ respectively. For example, wedging with $\mathbf{b}$, from the right, gives

$$x\,(\mathbf{a} \wedge \mathbf{b}) + y\,(\mathbf{b} \wedge \mathbf{b}) = \mathbf{c} \wedge \mathbf{b}, \tag{1.2}$$

but since $\mathbf{b} \wedge \mathbf{b} = 0$, we are left with

$$x\,(\mathbf{a} \wedge \mathbf{b}) = \mathbf{c} \wedge \mathbf{b}, \tag{1.3}$$

and since $\mathbf{a}, \mathbf{b}$ are not colinear, which means that $\mathbf{a} \wedge \mathbf{b} \neq 0$, we have

$$x = \frac{1}{\mathbf{a} \wedge \mathbf{b}} \mathbf{c} \wedge \mathbf{b}. \tag{1.4}$$

Similarly, we can wedge with $\mathbf{a}$ (from the left), to find

$$y = \frac{1}{\mathbf{a} \wedge \mathbf{b}} \mathbf{a} \wedge \mathbf{c}. \tag{1.5}$$

This works because, if the system has a solution, all the bivectors $\mathbf{a} \wedge \mathbf{b}$, $\mathbf{a} \wedge \mathbf{c}$, and $\mathbf{b} \wedge \mathbf{c}$, are all scalar multiples of each other, so we can just divide the two bivectors, and the results must be scalars.

1.2 Cramer's rule.

Incidentally, observe that for $\mathbb{R}^2$, this is the "Cramer's rule" solution to the system, since

$$\mathbf{x} \wedge \mathbf{y} = \begin{vmatrix} \mathbf{x} & \mathbf{y} \end{vmatrix} \mathbf{e}_1 \mathbf{e}_2, \tag{1.6}$$

where we are treating **x** and **y** here as column vectors of the coordinates. This means that, after dividing out the plane pseudoscalar $\mathbf{e}_1 \mathbf{e}_2$, we have

$$x = \frac{\begin{vmatrix} \mathbf{c} & \mathbf{b} \end{vmatrix}}{\begin{vmatrix} \mathbf{a} & \mathbf{b} \end{vmatrix}}$$

$$y = \frac{\begin{vmatrix} \mathbf{a} & \mathbf{c} \end{vmatrix}}{\begin{vmatrix} \mathbf{a} & \mathbf{b} \end{vmatrix}}. \tag{1.7}$$

This follows the usual Cramer's rule proscription, where we form determinants of the coordinates of the spanning vectors, replace either of the original vectors in the numerator with the target vector (depending on which variable we seek), and then take ratios of the two determinants.

1.3 Least squares solution, using geometry.

Now, let's consider the case, where the system eq. (1.1) cannot be solved exactly. Geometrically, the best we can do is to try to solve the related "least squares" problem

$$x\mathbf{a} + y\mathbf{b} = \mathbf{c}_{\parallel}, \tag{1.8}$$

where $\mathbf{c}_{\parallel}$ is the projection of **c** onto the plane spanned by $\mathbf{a}, \mathbf{b}$. Regardless of the value of **c**, we can always find a solution to this problem. For example, solving for x, we have

$$\begin{aligned} x &= \frac{1}{\mathbf{a} \wedge \mathbf{b}} \mathbf{c}_{\parallel} \wedge \mathbf{b} \\ &= \frac{1}{\mathbf{a} \wedge \mathbf{b}} \cdot \left(\mathbf{c}_{\parallel} \wedge \mathbf{b} \right) \\ &= \frac{1}{\mathbf{a} \wedge \mathbf{b}} \cdot \left(\mathbf{c} \wedge \mathbf{b} \right) - \frac{1}{\mathbf{a} \wedge \mathbf{b}} \cdot \left(\mathbf{c}_{\perp} \wedge \mathbf{b} \right). \end{aligned} \tag{1.9}$$

Let's look at the second term, which can be written

$$\begin{aligned} -\frac{1}{\mathbf{a} \wedge \mathbf{b}} \cdot \left(\mathbf{c}_{\perp} \wedge \mathbf{b} \right) &= -\frac{\mathbf{a} \wedge \mathbf{b}}{(\mathbf{a} \wedge \mathbf{b})^2} \cdot \left(\mathbf{c}_{\perp} \wedge \mathbf{b} \right) \\ &\propto (\mathbf{a} \wedge \mathbf{b}) \cdot \left(\mathbf{c}_{\perp} \wedge \mathbf{b} \right) \\ &= \left((\mathbf{a} \wedge \mathbf{b}) \cdot \mathbf{c}_{\perp} \right) \cdot \mathbf{b} \\ &= \left(\mathbf{a} \left(\mathbf{b} \cdot \mathbf{c}_{\perp} \right) - \mathbf{b} \left(\mathbf{a} \cdot \mathbf{c}_{\perp} \right) \right) \cdot \mathbf{b} \\ &= 0. \end{aligned} \tag{1.10}$$

The zero above follows because $\mathbf{c}_{\perp}$ is perpendicular to both **a** and **b** by construction. Geometrically, we are trying to dot two perpendicular bivectors, where **b** is a common factor of those two bivectors, as illustrated in fig. 1.1.

We see that our least squares solution, to this two variable linear system problem, is

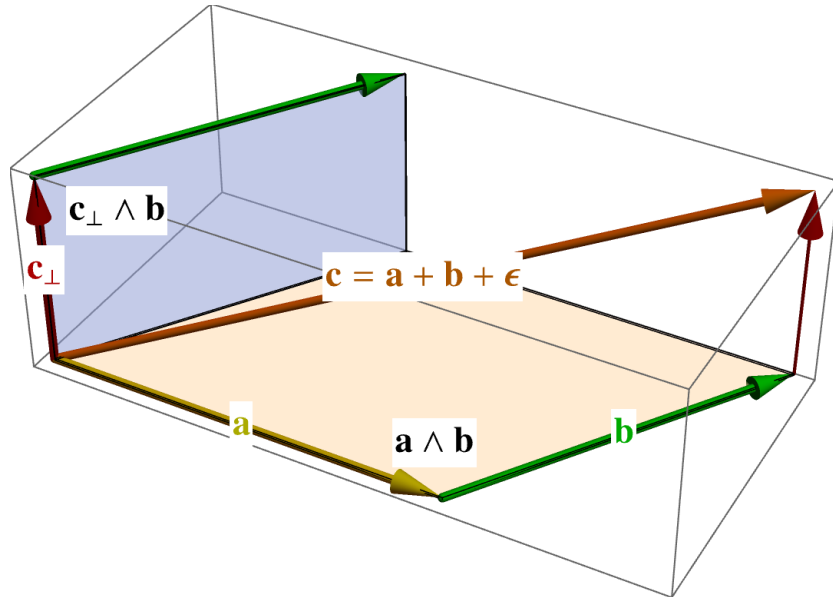$$x = \frac{1}{\mathbf{a} \wedge \mathbf{b}} \cdot \left(\mathbf{c} \wedge \mathbf{b} \right). \tag{1.11}$$

Figure 1.1: Perpendicular bivectors.

$$y = \frac{1}{\mathbf{a} \wedge \mathbf{b}} \cdot (\mathbf{a} \wedge \mathbf{c}).$$ (1.12)

The interesting thing here is how we have managed to connect the geometric notion of the optimal solution, the equivalent of a least squares solution (which we can compute with the Moore-Penrose inverse, or with an SVD (Singular Value Decomposition)), with the entirely geometric notion of selecting for the portion of the desired solution that lies within the span of the set of input vectors, provided that the spanning vectors for that hyperplane are linearly independent.

1.4   Least squares solution, using calculus.

I've called the projection solution, a least-squares solution, without full justification. Here's that justification. We define the usual error function, the squared distance from the target, from our superposition position in the plane

$$\epsilon = (\mathbf{c} - x\mathbf{a} - y\mathbf{b})^2,$$ (1.13)

and then take partials with respect to x, y, equating each to zero

$$0 = \frac{\partial \epsilon}{\partial x} = 2(\mathbf{c} - x\mathbf{a} - y\mathbf{b}) \cdot (-\mathbf{a})$$

$$0 = \frac{\partial \epsilon}{\partial y} = 2(\mathbf{c} - x\mathbf{a} - y\mathbf{b}) \cdot (-\mathbf{b}).$$ (1.14)

This is a two equation, two unknown system, which can be expressed in matrix form as

$$\begin{bmatrix} \mathbf{a}^2 & \mathbf{a} \cdot \mathbf{b} \\ \mathbf{a} \cdot \mathbf{b} & \mathbf{b}^2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \mathbf{a} \cdot \mathbf{c} \\ \mathbf{b} \cdot \mathbf{c} \end{bmatrix}. \tag{1.15}$$

This has solution

$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{\begin{vmatrix} \mathbf{a}^2 & \mathbf{a} \cdot \mathbf{b} \\ \mathbf{a} \cdot \mathbf{b} & \mathbf{b}^2 \end{vmatrix}} \begin{bmatrix} \mathbf{b}^2 & -\mathbf{a} \cdot \mathbf{b} \\ -\mathbf{a} \cdot \mathbf{b} & \mathbf{a}^2 \end{bmatrix} \begin{bmatrix} \mathbf{a} \cdot \mathbf{c} \\ \mathbf{b} \cdot \mathbf{c} \end{bmatrix} = \frac{\begin{bmatrix} \mathbf{b}^2 (\mathbf{a} \cdot \mathbf{c}) - (\mathbf{a} \cdot \mathbf{b}) (\mathbf{b} \cdot \mathbf{c}) \\ \mathbf{a}^2 (\mathbf{b} \cdot \mathbf{c}) - (\mathbf{a} \cdot \mathbf{b}) (\mathbf{a} \cdot \mathbf{c}) \end{bmatrix}}{\mathbf{a}^2 \mathbf{b}^2 - (\mathbf{a} \cdot \mathbf{b})^2}. \tag{1.16}$$

All of these differences can be expressed as wedge dot products, using the following expansions in reverse

$$\begin{aligned} (\mathbf{a} \wedge \mathbf{b}) \cdot (\mathbf{c} \wedge \mathbf{d}) &= \mathbf{a} \cdot (\mathbf{b} \cdot (\mathbf{c} \wedge \mathbf{d})) \\ &= \mathbf{a} \cdot ((\mathbf{b} \cdot \mathbf{c}) \mathbf{d} - (\mathbf{b} \cdot \mathbf{d}) \mathbf{c}) \\ &= (\mathbf{a} \cdot \mathbf{d}) (\mathbf{b} \cdot \mathbf{c}) - (\mathbf{a} \cdot \mathbf{c}) (\mathbf{b} \cdot \mathbf{d}). \end{aligned} \tag{1.17}$$

We find

$$\begin{aligned} x &= \frac{\mathbf{b}^2 (\mathbf{a} \cdot \mathbf{c}) - (\mathbf{a} \cdot \mathbf{b}) (\mathbf{b} \cdot \mathbf{c})}{-(\mathbf{a} \wedge \mathbf{b})^2} \\ &= \frac{(\mathbf{a} \wedge \mathbf{b}) \cdot (\mathbf{b} \wedge \mathbf{c})}{-(\mathbf{a} \wedge \mathbf{b})^2} \\ &= \frac{1}{\mathbf{a} \wedge \mathbf{b}} \cdot (\mathbf{c} \wedge \mathbf{b}), \end{aligned} \tag{1.18}$$

and

$$\begin{aligned} y &= \frac{\mathbf{a}^2 (\mathbf{b} \cdot \mathbf{c}) - (\mathbf{a} \cdot \mathbf{b}) (\mathbf{a} \cdot \mathbf{c})}{-(\mathbf{a} \wedge \mathbf{b})^2} \\ &= \frac{-(\mathbf{a} \wedge \mathbf{b}) \cdot (\mathbf{a} \wedge \mathbf{c})}{-(\mathbf{a} \wedge \mathbf{b})^2} \\ &= \frac{1}{\mathbf{a} \wedge \mathbf{b}} \cdot (\mathbf{a} \wedge \mathbf{c}). \end{aligned} \tag{1.19}$$

Sure enough, we find what was dubbed our least squares solution, which we now know can be written out as a ratio of (dotted) wedge products. From eq. (1.15), it wasn't obvious that the least squares solution would have a structure that was almost Cramer's rule like, but having solved this problem using geometry alone, we knew to expect that. It was therefore natural to write the results in terms of wedge products factors, and find the simplest statement of the end result. That end result reduces to Cramer's rule for the $\mathbb{R}^2$ special case where the system has an exact solution.