Beginner's guide to mapping simplexes affinely

Preprint · April 2019

DOI: 10.13140/RG.2.2.13787.41762

2 authors:

Vitalii Bohdanovych Tymchyshyn
National Academy of Sciences of Ukraine

Andrii Khlevniuk

Beginner's guide to mapping simplexes affinely

V. B. Tymchyshyn*[,1] and A. V. Khlevniuk

[1]Bogolyubov Institute for Theoretical Physics, National Academy of Sciences, Metrolohichna St. 14-b, Kyiv 03680, Ukraine

August 11, 2019

Abstract

What is the affine transformation that maps 3 points in a plane into another 3 points? Can I map affinely any 4 points in 3D into another arbitrary 4 points? What about n-dimensional space and why does anyone care about these transformations? If you feel interested in any of these questions, this review will provide you the answers.

Contents

*corresponding author, yu.binkukoku@gmail.com

Intuitive introduction

Sometimes we want to establish a mapping of space being guided "by examples" — knowing images of certain points we demand all the rest points to "follow the same rule." Surprisingly, there are much more examples of this problem than one may expect.

Consider texturing — a part of image should be applied (mapped) to a polygon (often a triangle). In most cases we know which pixels of the image correspond to vertices of the polygon, but mapping for internal points should be deduced. Thus the problem — how do we convert our knowledge about vertices mapping into transformation for the entire space? Note, it may include shears in addition to rotations, translations and scale.

Consider a different problem. Say we create an augmented reality application that detects a hand and places a virtual kitten on it. Our recognition engine periodically reports 3D-coordinates of the points of the hand. How should we set the coordinates of the kitten, so that it stays on the hand all the time? It's not so easy if you think about possible rotations. And once again, we have the problem of mapping points "the same way" certain other points were mapped.

Somewhat unexpectedly, multilinear interpolation is an affine mapping problem as well, but this time the role of codomain plays the "property space", e.g. space of RGB colors. If we define some colors for vertices of the triangle or pyramid, how do we perform the multilinear interpolation across its interior? This problem has a "mapping flavor" as it seems to be transferring our simplex to "color space" thus every point acquires color. Very similar problem appears for Phong shading — we define vectors at certain points and want to perform the multilinear interpolation.

Now as we've gained intuition of what we want, we are ready to formulate the problem mathematically.

Problem statement

What kind of mapping are we looking for?

There are many ways to map Euclidean space to itself and specifying images of a few points is not very helpful — the rest of the points can still be mapped virtually anywhere. We need to narrow the class of space transformations to a meaningful subclass, such that its member can be unambiguously determined by the way it transforms a small number of points.

As such it's reasonable to consider affine transformations. They include translation (i.e. shift), rotation, scale, and reflection — the most common space transformations. Moreover, as we show below, by specifying the images of just a handful of points we will be able to unambiguously recover the transformation itself.

Let's recap what affine transformation is. Affine transformation in Euclidean space is described as follows

$$\vec{X}(\vec{x}) = \hat{A}\vec{x} + \vec{t}, \tag{1}$$

where $\vec{x}$ is the point we are mapping, $\vec{X}(\vec{x})$ is its image, $\hat{A}$ — transformation matrix, and $\vec{t}$ — translation vector. Often $\hat{A}$ and $\vec{t}$ are "glued together" into an augmented matrix so that the whole transformation is described with the single matrix of certain form instead of pair matrix-vector.

How many points do we need?

Once we are sure, it's the affine transformation we are looking for, we are facing another question — how many points and their images are needed to uniquely pinpoint the exact mapping?

We start with 2D space for simplicity and consider (1). Suppose, a point $\vec{x}^{(1)}$ is mapped into $\vec{X}^{(1)}$. Upper index "(1)" is just a designation that this is the first point (and a hint there will be more). Does one point suffice to determine a 2D affine transformation? We write (1) for this point and get following

$$\vec{X}^{(1)} = \hat{A}\vec{x}^{(1)} + \vec{t}, \Leftrightarrow \begin{pmatrix} X_1^{(1)} \\ X_2^{(1)} \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} \begin{pmatrix} x_1^{(1)} \\ x_2^{(1)} \end{pmatrix} + \begin{pmatrix} t_1 \\ t_2 \end{pmatrix}, \Leftrightarrow \begin{cases} X_1^{(1)} = a_{1,1}x_1^{(1)} + a_{1,2}x_2^{(1)} + t_1, \\ X_2^{(1)} = a_{2,1}x_1^{(1)} + a_{2,2}x_2^{(1)} + t_2, \end{cases}$$

where subindices at x and X designate coordinate, a-s are elements of matrix $\hat{A}$, and t-s are components of translation vector $\vec{t}$. The answer to our question is negative — we have 6 parameters to determine (four a-s and two t-s) with only 2 linear equations given.

It becomes obvious, we need at least 6 equations, thus 3 points and their images should be available

$$\underbrace{\begin{aligned} X_1^{(1)} &= a_{1,1}x_1^{(1)} + a_{1,2}x_2^{(1)} + t_1, \\ X_2^{(1)} &= a_{2,1}x_1^{(1)} + a_{2,2}x_2^{(1)} + t_2, \end{aligned}}_{\vec{x}^{\,(1)} \to \vec{X}^{(1)}} \quad \underbrace{\begin{aligned} X_1^{(2)} &= a_{1,1}x_1^{(2)} + a_{1,2}x_2^{(2)} + t_1, \\ X_2^{(2)} &= a_{2,1}x_1^{(2)} + a_{2,2}x_2^{(2)} + t_2, \end{aligned}}_{\vec{x}^{\,(2)} \to \vec{X}^{(2)}} \quad \underbrace{\begin{aligned} X_1^{(3)} &= a_{1,1}x_1^{(3)} + a_{1,2}x_2^{(3)} + t_1, \\ X_2^{(3)} &= a_{2,1}x_1^{(3)} + a_{2,2}x_2^{(3)} + t_2. \end{aligned}}_{\vec{x}^{\,(3)} \to \vec{X}^{(3)}}$$

These equations can be merged together and written in a matrix form, so that we can clearly see the parameters given as well as the unknowns

$$\begin{pmatrix} X_1^{(1)} \\ X_1^{(2)} \\ X_1^{(3)} \\ X_2^{(1)} \\ X_2^{(2)} \\ X_2^{(3)} \end{pmatrix} = \begin{pmatrix} x_1^{(1)} & x_2^{(1)} & 1 & 0 & 0 & 0 \\ x_1^{(2)} & x_2^{(2)} & 1 & 0 & 0 & 0 \\ x_1^{(3)} & x_2^{(3)} & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_1^{(1)} & x_2^{(1)} & 1 \\ 0 & 0 & 0 & x_1^{(2)} & x_2^{(2)} & 1 \\ 0 & 0 & 0 & x_1^{(3)} & x_2^{(3)} & 1 \end{pmatrix} \begin{pmatrix} a_{1,1} \\ a_{1,2} \\ t_1 \\ a_{2,1} \\ a_{2,2} \\ t_2 \end{pmatrix}.$$

The system has a unique solution if and only if the determinant of the matrix is non-zero. Since we are dealing with block matrix, it suffices to demand the block

$$\hat{B}_{[2]} = \begin{pmatrix} x_1^{(1)} & x_2^{(1)} & 1 \\ x_1^{(2)} & x_2^{(2)} & 1 \\ x_1^{(3)} & x_2^{(3)} & 1 \end{pmatrix}$$

to have non-zero determinant. Index "[2]" means the number of dimensions (we consider 2D), later we will consider another spaces.

When does $B_{[2]}$ have non-zero determinant? We choose any row of $B_{[2]}$, subtract it from all the rest rows, and then use Laplace expansion along the last column. As a result we get

$$\det\left(\hat{B}_{[2]}\right) = \det\begin{pmatrix} x_1^{(2)} - x_1^{(1)} & x_2^{(2)} - x_2^{(1)} \\ x_1^{(3)} - x_1^{(1)} & x_2^{(3)} - x_2^{(1)} \end{pmatrix},$$

which means, determinant of $B_{[2]}$ is not equal to zero if and only if vectors $\vec{x}^{\,(2)} - \vec{x}^{\,(1)}$ and $\vec{x}^{\,(3)} - \vec{x}^{\,(1)}$ are linearly independent, i.e. points $\vec{x}^{\,(i)}$ form a *simplex* — non-degenerate triangle in this case.

We can reproduce all the considerations for higher dimensions and find out, that affine transformation can be uniquely determined if and only if the block

$$\hat{B}_{[n]} = \begin{pmatrix} x_1^{(1)} & x_2^{(1)} & \cdots & x_n^{(1)} & 1 \\ x_1^{(2)} & x_2^{(2)} & \cdots & x_n^{(2)} & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_1^{(n+1)} & x_2^{(n+1)} & \cdots & x_n^{(n+1)} & 1 \end{pmatrix}$$

has non-zero determinant, thus differences $\vec{x}^{\,(i)} - \vec{x}^{\,(1)}$ for $i \neq 1$ are linearly independent, i.e. points $\vec{x}^{\,(i)}$ and point $\vec{x}^{\,(1)}$ form a simplex. Note, we do not impose any restrictions on images $\vec{X}^{(i)}$, thus transformations we consider include projections.

In conclusion, **affine transformation with n-dimensional linear space as its domain can be unambiguously defined by its action on vertices of n-dimensional simplex ($n + 1$ points)**. Now we are ready for the

Rigorous problem statement

Find the affine transformation that maps vertices of simplex $\vec{x}^{\,(1)}$,..., $\vec{x}^{\,(n+1)}$ into points $\vec{X}^{(1)}, \ldots, \vec{X}^{(n+1)}$.

Solution

Assume n-dimensional simplex S has vertices $\vec{x}^{(1)}, \dots, \vec{x}^{(n+1)}$. Affine transformation that maps them into $\vec{X}^{(1)}, \dots, \vec{X}^{(n+1)}$ has the following form

$$\vec{X}(\vec{x}) = (-1)\frac{\det\begin{pmatrix} 0 & \vec{X}^{(1)} & \vec{X}^{(2)} & \dots & \vec{X}^{(n+1)} \\ x_1 & x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(n+1)} \\ x_2 & x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(n+1)} \\ \dots & \dots & \dots & \dots & \dots \\ x_n & x_n^{(1)} & x_n^{(2)} & \dots & x_n^{(n+1)} \\ 1 & 1 & 1 & \dots & 1 \end{pmatrix}}{\det\begin{pmatrix} x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(n+1)} \\ x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(n+1)} \\ \dots & \dots & \dots & \dots \\ x_n^{(1)} & x_n^{(2)} & \dots & x_n^{(n+1)} \\ 1 & 1 & \dots & 1 \end{pmatrix}}. \tag{2}$$

Well... that's all. For the rest of the article we'll be busy exploring this formula.

Structure of the matrix

Matrices in (2) have a neat structure.

$$\vec{X}(\vec{x}) = (-1)\frac{\det\begin{pmatrix} 0 & \vec{X}^{(1)} & \vec{X}^{(2)} & \dots & \vec{X}^{(n+1)} \\ \boxed{\begin{matrix} x_1 \\ x_2 \\ \dots \\ x_n \end{matrix}} & \boxed{\begin{matrix} x_1^{(1)} \\ x_2^{(1)} \\ \dots \\ x_n^{(1)} \end{matrix}} & \boxed{\begin{matrix} x_1^{(2)} \\ x_2^{(2)} \\ \dots \\ x_n^{(2)} \end{matrix}} & \dots & \boxed{\begin{matrix} x_1^{(n+1)} \\ x_2^{(n+1)} \\ \dots \\ x_n^{(n+1)} \end{matrix}} \\ 1 & 1 & 1 & \dots & 1 \end{pmatrix}}{\det\begin{pmatrix} \boxed{\begin{matrix} x_1^{(1)} \\ x_2^{(1)} \\ \dots \\ x_n^{(1)} \end{matrix}} & \boxed{\begin{matrix} x_1^{(2)} \\ x_2^{(2)} \\ \dots \\ x_n^{(2)} \end{matrix}} & \dots & \boxed{\begin{matrix} x_1^{(n+1)} \\ x_2^{(n+1)} \\ \dots \\ x_n^{(n+1)} \end{matrix}} \\ 1 & 1 & \dots & 1 \end{pmatrix}}. \tag{3}$$

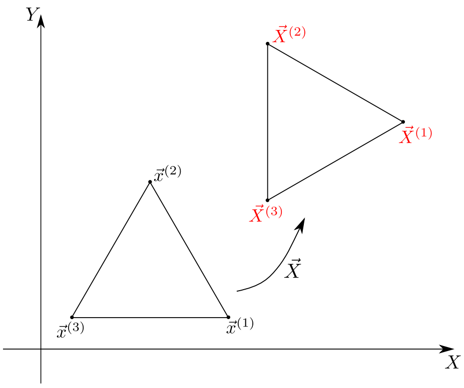
Matrix in the numerator (3) is like a hamburger — vector-columns (in boxes) are positioned between row of ones from below and row of vectors $\vec{X}^{(i)}$ from above (except for the left-most entry). The left-most vector-column contains arguments of the function $\vec{X}$, i.e. coordinates of the vector argument of the function. It's easily distinguishable since components of $\vec{x}$ have no bracketed upper index (in (3) color-coded with blue). Above this vector-column there is a zero (what else could we expect on the intersection of numbers' column and vectors' row?). All the rest vector-columns are parameters of the affine transformation — coordinates of the points $\vec{x}^{(i)}$ that constitute the simplex. Above every vector-column $\vec{x}^{(i)}$ there's a vector $\vec{X}^{(i)}$ that $\vec{x}^{(i)}$ is mapped to (in (3) color-coded with red). The bottom row consists of just ones.

Note that vector $\vec{X}^{(i)}$ is taken as a single entity, not as a collection of components. You have already seen this notation before — cross product of two 3D vectors is often written as a determinant of a matrix with standard basis vectors $\vec{i}, \vec{j}, \vec{k}$ in the first row. You should think about (2) the same way.

Matrix in the denominator (2) is just the matrix in the numerator with the left-most column and the top-most row dropped.

Low-dimensional example

Before we proceed, consider a 2D example to see how the things work. Image below shows three points $\vec{x}^{(1)}$, $\vec{x}^{(2)}$, and $\vec{x}^{(3)}$ that are affinely transformed into $\vec{X}^{(1)}$, $\vec{X}^{(2)}$, and $\vec{X}^{(3)}$.



As usual $\vec{i}$ and $\vec{j}$ are standard unit vectors. Let points $\vec{x}^{\,(1)}$, $\vec{x}^{\,(2)}$, and $\vec{x}^{\,(3)}$ have the following coordinates

$$\begin{aligned}\vec{x}^{\,(1)} &= \vec{i},\\ \vec{x}^{\,(2)} &= \sqrt{3}\,\vec{j},\\ \vec{x}^{\,(3)} &= -\vec{i}.\end{aligned}$$

Under transformation $\vec{X}$ they turn into

$$\begin{aligned}\vec{X}(\vec{x}^{\,(1)}) &= \vec{X}^{\,(1)} = \frac{3+2\sqrt{3}}{2}\vec{i} + \frac{5}{2}\vec{j},\\ \vec{X}(\vec{x}^{\,(2)}) &= \vec{X}^{\,(2)} = \frac{3}{2}\vec{i} + \frac{7}{2}\vec{j},\\ \vec{X}(\vec{x}^{\,(3)}) &= \vec{X}^{\,(3)} = \frac{3}{2}\vec{i} + \frac{3}{2}\vec{j}.\end{aligned}$$

We want to find the affine transform $\vec{X}$. Let's start with general expression (2) for 2D space (we keep color-coding as in (3))

$$\vec{X}(\vec{x}\,) = (-1)\frac{\det\begin{pmatrix} 0 & \vec{X}^{(1)} & \vec{X}^{(2)} & \vec{X}^{(3)} \\ x_1 & x_1^{(1)} & x_1^{(2)} & x_1^{(3)} \\ x_2 & x_2^{(1)} & x_2^{(2)} & x_2^{(3)} \\ 1 & 1 & 1 & 1 \end{pmatrix}}{\det\begin{pmatrix} x_1^{(1)} & x_1^{(2)} & x_1^{(3)} \\ x_2^{(1)} & x_2^{(2)} & x_2^{(3)} \\ 1 & 1 & 1 \end{pmatrix}} = \frac{-\det\begin{pmatrix} 0 & \frac{3+2\sqrt{3}}{2}\vec{i} + \frac{5}{2}\vec{j} & \frac{3}{2}\vec{i} + \frac{7}{2}\vec{j} & \frac{3}{2}\vec{i} + \frac{3}{2}\vec{j} \\ x_1 & 1 & 0 & -1 \\ x_2 & 0 & \sqrt{3} & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}}{\det\begin{pmatrix} 1 & 0 & -1 \\ 0 & \sqrt{3} & 0 \\ 1 & 1 & 1 \end{pmatrix}},$$

after simplification that becomes

$$\vec{X}(\vec{x}\,) = \frac{\sqrt{3}\,\vec{i} + \vec{j}}{2}x_1 - \frac{\vec{i} - \sqrt{3}\,\vec{j}}{2}x_2 + \frac{3+\sqrt{3}}{2}\vec{i} + 2\vec{j} = \begin{pmatrix} \sqrt{3}/2 & -1/2 \\ 1/2 & \sqrt{3}/2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} (3+\sqrt{3})/2 \\ 2 \end{pmatrix}.$$

Now one can clearly spot 30° rotation and translation as we resorted to canonical notation. It is worth noting, expression (2) is convenient for mapping points with rational coordinates to points with rational coordinates — calculations are simple, no transcendental numbers or functions needed.

Deriving formula for 3D case is left to the Reader as an exercise.

Properties

The function is well-defined

There are two issues about (2): can denominator be zero and how to calculate determinant of the matrix in the numerator, since it contains vectors as elements.

The first issue is easy to resolve. Let's move origin to the vertex $\vec{x}^{\,(1)}$. Since S is a simplex, we know that all $\vec{x}^{\,(i)} - \vec{x}^{\,(1)}$ for $i \neq 1$ are linearly independent. Subtract first column in a matrix in the denominator (2) from the rest of the columns and use the Laplace expansion formula along the last row

$$\det\begin{pmatrix} x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(n+1)} \\ x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(n+1)} \\ \dots & \dots & \dots & \dots \\ x_n^{(1)} & x_n^{(2)} & \dots & x_n^{(n+1)} \\ 1 & 1 & \dots & 1 \end{pmatrix} = \det\begin{pmatrix} x_1^{(2)} - x_1^{(1)} & \dots & x_1^{(n+1)} - x_1^{(1)} \\ x_2^{(2)} - x_2^{(1)} & \dots & x_2^{(n+1)} - x_2^{(1)} \\ \dots & \dots & \dots \\ x_n^{(2)} - x_n^{(1)} & \dots & x_n^{(n+1)} - x_n^{(1)} \end{pmatrix} \neq 0,$$

the inequality appears since all columns are linearly independent (S is a simplex).

As for the determinant in the numerator (2) it is easy to show that it's well-formed. We expand it along the first row — the one with $\vec{X}^{(i)}$-s

$$\det\begin{pmatrix} 0 & \vec{X}^{(1)} & \vec{X}^{(2)} & \dots & \vec{X}^{(n+1)} \\ x_1 & x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(n+1)} \\ x_2 & x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(n+1)} \\ \dots & \dots & \dots & \dots & \dots \\ x_n & x_n^{(1)} & x_n^{(2)} & \dots & x_n^{(n+1)} \\ 1 & 1 & 1 & \dots & 1 \end{pmatrix} = (-1)^1\vec{X}^{(1)}\det\begin{pmatrix} x_1 & x_1^{(2)} & \dots & x_1^{(n+1)} \\ x_2 & x_2^{(2)} & \dots & x_2^{(n+1)} \\ \dots & \dots & \dots & \dots \\ x_n & x_n^{(2)} & \dots & x_n^{(n+1)} \\ 1 & 1 & \dots & 1 \end{pmatrix} + \cdots$$

$$\cdots + (-1)^{n+1}\vec{X}^{(n+1)}\det\begin{pmatrix} x_1 & x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(n)} \\ x_2 & x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(n)} \\ \dots & \dots & \dots & \dots & \dots \\ x_n & x_n^{(1)} & x_n^{(2)} & \dots & x_n^{(n)} \\ 1 & 1 & 1 & \dots & 1 \end{pmatrix}.$$

Matrices in the last equation contain only real numbers, thus their determinants are well-defined. The whole expression is a linear combination of vectors $\vec{X}^{(i)}$-s that results an ordinary vector.

The function is affine

Note, the determinant in the denominator (2) contains scalars and that's why equals a scalar itself. The numerator (2) contains linear combination of variables x_i-s plus a constant — result easily obtained from the Laplace expansion formula along the first column

$$\det\begin{pmatrix} 0 & \vec{X}^{(1)} & \vec{X}^{(2)} & \dots & \vec{X}^{(n+1)} \\ x_1 & x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(n+1)} \\ x_2 & x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(n+1)} \\ \dots & \dots & \dots & \dots & \dots \\ x_n & x_n^{(1)} & x_n^{(2)} & \dots & x_n^{(n+1)} \\ 1 & 1 & 1 & \dots & 1 \end{pmatrix} = (-1)^1 x_1\det\begin{pmatrix} \vec{X}^{(1)} & \vec{X}^{(2)} & \dots & \vec{X}^{(n+1)} \\ x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(n+1)} \\ \dots & \dots & \dots & \dots \\ x_n^{(1)} & x_n^{(2)} & \dots & x_n^{(n+1)} \\ 1 & 1 & \dots & 1 \end{pmatrix} + \cdots$$

$$\cdots + (-1)^n x_n\det\begin{pmatrix} \vec{X}^{(1)} & \vec{X}^{(2)} & \dots & \vec{X}^{(n+1)} \\ x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(n+1)} \\ x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(n+1)} \\ \dots & \dots & \dots & \dots \\ x_{n-1}^{(1)} & x_{n-1}^{(2)} & \dots & x_{n-1}^{(n+1)} \\ 1 & 1 & \dots & 1 \end{pmatrix} + (-1)^{n+1}\det\begin{pmatrix} \vec{X}^{(1)} & \vec{X}^{(2)} & \dots & \vec{X}^{(n+1)} \\ x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(n+1)} \\ x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(n+1)} \\ \dots & \dots & \dots & \dots \\ x_n^{(1)} & x_n^{(2)} & \dots & x_n^{(n+1)} \end{pmatrix},$$

which means function $\vec{X}$ is affine.

The function maps $\vec{x}^{(i)}$ into $\vec{X}^{(i)}$

That's the most interesting property to prove. We would like to show that $\vec{X}(\vec{x}^{(i)}) = \vec{X}^{(i)}$. Since full formula is cumbersome, first we consider the numerator of (2). If we substitute $\vec{x}^{(i)}$ for $\vec{x}$

$$\det\begin{pmatrix} 0 & \vec{X}^{(1)} & \dots & \vec{X}^{(i)} & \dots & \vec{X}^{(n+1)} \\ x_1^{(i)} & x_1^{(1)} & \dots & x_1^{(i)} & \dots & x_1^{(n+1)} \\ x_2^{(i)} & x_2^{(1)} & \dots & x_2^{(i)} & \dots & x_2^{(n+1)} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_n^{(i)} & x_n^{(1)} & \dots & x_n^{(i)} & \dots & x_n^{(n+1)} \\ 1 & 1 & \dots & 1 & \dots & 1 \end{pmatrix} = \det\begin{pmatrix} -\vec{X}^{(i)} & \vec{X}^{(1)} & \dots & \vec{X}^{(i)} & \dots & \vec{X}^{(n+1)} \\ 0 & x_1^{(1)} & \dots & x_1^{(i)} & \dots & x_1^{(n+1)} \\ 0 & x_2^{(1)} & \dots & x_2^{(i)} & \dots & x_2^{(n+1)} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & x_n^{(1)} & \dots & x_n^{(i)} & \dots & x_n^{(n+1)} \\ 0 & 1 & \dots & 1 & \dots & 1 \end{pmatrix} =$$

$$= -\vec{X}^{(i)}\det\begin{pmatrix} x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(n+1)} \\ x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(n+1)} \\ \dots & \dots & \dots & \dots \\ x_n^{(1)} & x_n^{(2)} & \dots & x_n^{(n+1)} \\ 1 & 1 & \dots & 1 \end{pmatrix}.$$

Here we substituted $\vec{x}^{\,(i)}$ for $\vec{x}$, found the column with $\vec{X}^{(i)}$, and subtracted it from the first column. Then we used the Laplace expansion formula along the first column and obtained simplified expression for the numerator of (2).

Now we can substitute the expression above into (2)

$$\vec{X}(\vec{x}^{\,(i)}) = (-1)\frac{-\vec{X}^{(i)}\det\begin{pmatrix} x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(n+1)} \\ x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(n+1)} \\ \dots & \dots & \dots & \dots \\ x_n^{(1)} & x_n^{(2)} & \dots & x_n^{(n+1)} \\ 1 & 1 & \dots & 1 \end{pmatrix}}{\det\begin{pmatrix} x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(n+1)} \\ x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(n+1)} \\ \dots & \dots & \dots & \dots \\ x_n^{(1)} & x_n^{(2)} & \dots & x_n^{(n+1)} \\ 1 & 1 & \dots & 1 \end{pmatrix}} = \vec{X}^{(i)},$$

that completes the proof.

Any affine transformation can be represented this way

This statement is obvious since we can take vertices of any simplex, map it by given affine transformation and plug all this into (2). Due to the previous statements, we will recover the initial affine mapping.

Fixed points

Suppose we are mapping from n-dimensional space to its subspace (not necessarily proper). If we set equal $\vec{X}^{\,(i)}$ to certain $\vec{x}^{(i)}$ — (2) will do the rest by providing us with transformation that maps $\vec{x}^{\,(i)}$ into itself.

Canonical notation

Recall we defined affine transformation (1)

$$\vec{X}(\vec{x}) = \hat{A}\vec{x} + \vec{t}.$$

We can easily get $\hat{A}$ and $\vec{t}$ from (2) using Laplace expansion along the first column

$$\hat{A} = -\frac{1}{C_{1,1}}\begin{pmatrix} \boxed{C_{2,1}} & \boxed{C_{3,1}} & \cdots & \boxed{C_{i+1,1}} & \cdots & \boxed{C_{n+1,1}} \end{pmatrix}; \qquad \vec{t} = -\frac{1}{C_{1,1}}\boxed{C_{n+2,1}},$$

where $C_{i,j}$ is a cofactor of the matrix

$$\begin{pmatrix} 0 & \vec{X}^{(1)} & \vec{X}^{(2)} & \dots & \vec{X}^{(n+1)} \\ x_1 & x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(n+1)} \\ x_2 & x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(n+1)} \\ \dots & \dots & \dots & \dots & \dots \\ x_n & x_n^{(1)} & x_n^{(2)} & \dots & x_n^{(n+1)} \\ 1 & 1 & 1 & \dots & 1 \end{pmatrix}.$$

Note in $C_{i,j}$ first index designates row and the second — column, thus $C_{i+1,1}$ corresponds to minor obtained by cutting of the first column and the $i+1$-*th* row (the one containing x_i). Since all vectors in the matrix under consideration are contained within the first row, the cofactor $C_{1,1}$ is a scalar, while all the rest cofactors $C_{i>1,1}$ are vector-columns— stacked together $C_{2,1},\ldots,C_{n+1,1}$ form matrix $\hat{A}$, while $C_{n+2,1}$ defines translation vector $\vec{t}$.

We can also construct (2) from $\hat{A}$ and $\vec{t}$. If

$$\hat{A}=(\vec{a_1},\vec{a_2},\ldots,\vec{a_n}),$$

where $\vec{a}_i$*'s* are vector columns we can write

$$\vec{X}(\vec{x})=(-1)\det\begin{pmatrix} 0 & \vec{a}_1+\vec{t} & \vec{a}_2+\vec{t} & \ldots & \vec{a}_n+\vec{t} & \vec{t} \\ x_1 & 1 & 0 & \ldots & 0 & 0 \\ x_2 & 0 & 1 & \ldots & 0 & 0 \\ \multicolumn{6}{c}{\dotfill} \\ x_n & 0 & 0 & \ldots & 1 & 0 \\ 1 & 1 & 1 & \ldots & 1 & 1 \end{pmatrix}.$$

as transformation $\hat{A}\vec{x}+\vec{t}$ sends simplex $\{\vec{e}_1,\vec{e}_2\ldots,\vec{e}_n,\vec{0}\,\}$ to $\{\vec{a}_1+\vec{t},\vec{a}_2+\vec{t}\ldots,\vec{a}_n+\vec{t},\vec{t}\,\}$. If $\hat{A}$ allows for inverse transformation, it can be easily written as well just by exchanging $\vec{e}_i$-*s* and $\vec{a}_i+\vec{t}$-*s* places, i.e. acting with direct transformation we gain knowledge on point-image pairs and then swap domain with codomain.

Note, that canonical notation is more convenient when affine transformation matches simplex $\{\vec{e}_1$, $\vec{e}_2$, $\ldots$, $\vec{e}_n$, $\vec{0}\,\}$ into arbitrary simplex, while (2) is easier to write when simplex in general position is mapped into another simplex in general position.

Connection to other problems

Barycentric coordinates

For the sake of simplicity we consider 2D case. Suppose, we affinely transform triangle with vertices $\vec{x}^{(1)}$, $\vec{x}^{(2)}$, and $\vec{x}^{(3)}$ to triangle with vertices $\vec{X}^{(1)}$, $\vec{X}^{(2)}$, and $\vec{X}^{(3)}$. Affine transformations are known to be preserving barycentric coordinates of the points. Thus, barycentric coordinates $\{\lambda_1,\lambda_2,\lambda_3\}$ of any point should be equal before and after affine transformation

$$\begin{aligned}\vec{x}&=\lambda_1\vec{x}^{(1)}+\lambda_2\vec{x}^{(2)}+\lambda_3\vec{x}^{(3)},\\ \vec{X}&=\lambda_1\vec{X}^{(1)}+\lambda_2\vec{X}^{(2)}+\lambda_3\vec{X}^{(3)},\\ 1&=\lambda_1+\lambda_2+\lambda_3,\end{aligned}$$

where the last equality is the generic property of barycentric coordinates.

We can use the first and the last equations from the previous triple and write them in a component-wise form

$$\begin{aligned}x_1&=\lambda_1x_1^{(1)}+\lambda_2x_1^{(2)}+\lambda_3x_1^{(3)},\\ x_2&=\lambda_1x_2^{(1)}+\lambda_2x_2^{(2)}+\lambda_3x_2^{(3)},\\ 1&=\lambda_1+\lambda_2+\lambda_3.\end{aligned}$$

The Cramer's rule allows us to write solution for barycentric coordinates as follows

$$\lambda_1=\frac{1}{\Delta}\det\begin{pmatrix} x_1 & x_1^{(2)} & x_1^{(3)} \\ x_2 & x_2^{(2)} & x_2^{(3)} \\ 1 & 1 & 1\end{pmatrix},\ \lambda_2=\frac{1}{\Delta}\det\begin{pmatrix} x_1^{(1)} & x_1 & x_1^{(3)} \\ x_2^{(1)} & x_2 & x_2^{(3)} \\ 1 & 1 & 1\end{pmatrix},\ \lambda_3=\frac{1}{\Delta}\det\begin{pmatrix} x_1^{(1)} & x_1^{(2)} & x_1 \\ x_2^{(1)} & x_2^{(2)} & x_2 \\ 1 & 1 & 1\end{pmatrix},$$

where

$$\Delta=\det\begin{pmatrix} x_1^{(1)} & x_1^{(2)} & x_1^{(3)} \\ x_2^{(1)} & x_2^{(2)} & x_2^{(3)} \\ 1 & 1 & 1\end{pmatrix}.$$

Now plug λ-s into the equation for $\vec{X}$ and rearrange columns in the determinants

$$\vec{X} = \frac{\vec{X}^{(1)} \det \begin{pmatrix} x_1 & x_1^{(2)} & x_1^{(3)} \\ x_2 & x_2^{(2)} & x_2^{(3)} \\ 1 & 1 & 1 \end{pmatrix} - \vec{X}^{(2)} \det \begin{pmatrix} x_1 & x_1^{(1)} & x_1^{(3)} \\ x_2 & x_2^{(1)} & x_2^{(3)} \\ 1 & 1 & 1 \end{pmatrix} + \vec{X}^{(3)} \det \begin{pmatrix} x_1 & x_1^{(1)} & x_1^{(2)} \\ x_2 & x_2^{(1)} & x_2^{(2)} \\ 1 & 1 & 1 \end{pmatrix}}{\det \begin{pmatrix} x_1^{(1)} & x_1^{(2)} & x_1^{(3)} \\ x_2^{(1)} & x_2^{(2)} & x_2^{(3)} \\ 1 & 1 & 1 \end{pmatrix}}.$$

The last equation clearly looks like the Laplace expansion along the first row, i.e. we can write

$$\vec{X} = (-1)\frac{\det \begin{pmatrix} 0 & \vec{X}^{(1)} & \vec{X}^{(2)} & \vec{X}^{(3)} \\ x_1 & x_1^{(1)} & x_1^{(2)} & x_1^{(3)} \\ x_2 & x_2^{(1)} & x_2^{(2)} & x_2^{(3)} \\ 1 & 1 & 1 & 1 \end{pmatrix}}{\det \begin{pmatrix} x_1^{(1)} & x_1^{(2)} & x_1^{(3)} \\ x_2^{(1)} & x_2^{(2)} & x_2^{(3)} \\ 1 & 1 & 1 \end{pmatrix}} = \lambda_1 \vec{X}^{(1)} + \lambda_2 \vec{X}^{(2)} + \lambda_3 \vec{X}^{(3)},$$

that coincides with (2) for 2D case (Reader may also check the "Low-dimensional example" section). In a similar way it can be extended to any number of dimensions.

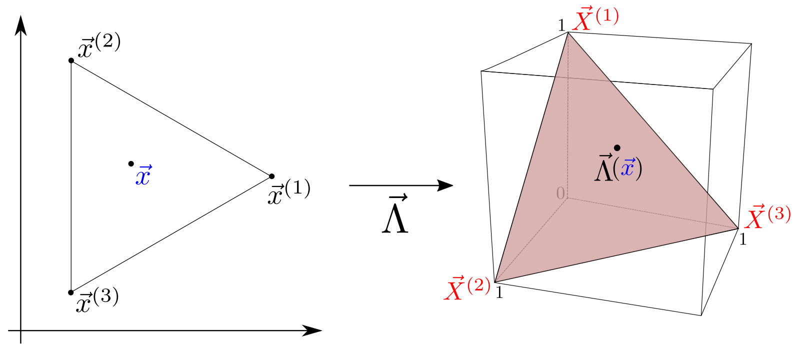


Figure 1: Calculation of barycentric coordinates

Worth noting, that (2) can be used the other way around — we can obtain barycentric coordinates of the point $\vec{x}$ with regard to the given simplex S by replacing (formally) $\vec{X}^{(i)}$-s with some orthonormal vectors $\vec{e}_i$-s and doing the determinants. As a result we will get vector in some $n+1$-dimensional space, whose components are barycentric coordinates of the point

$$\vec{\Lambda}(\vec{x}\,) = (-1)\frac{\det \begin{pmatrix} 0 & \vec{e}_1 & \vec{e}_2 & \dots & \vec{e}_{n+1} \\ x_1 & x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(n+1)} \\ x_2 & x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(n+1)} \\ \dots & \dots & \dots & \dots & \dots \\ x_n & x_n^{(1)} & x_n^{(2)} & \dots & x_n^{(n+1)} \\ 1 & 1 & 1 & \dots & 1 \end{pmatrix}}{\det \begin{pmatrix} x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(n+1)} \\ x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(n+1)} \\ \dots & \dots & \dots & \dots \\ x_n^{(1)} & x_n^{(2)} & \dots & x_n^{(n+1)} \\ 1 & 1 & \dots & 1 \end{pmatrix}} = \sum_{k=1}^{n+1} \vec{e}_k \lambda_k = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \dots \\ \lambda_{n+1} \end{pmatrix}. \qquad (4)$$

The last equation may be useful for finding points of curve-simplex intersection. Besides this form shows that calculation of barycentric coordinates is an affine transformation that maps n-dimensional simplex to $n+1$ dimensional space so that it becomes the standard n-dimensional simplex (i.e. its vertices become "tips" of orthonormal coordinate vectors) as shown in figure 1.

Homogeneous coordinates

We have already proven that the determinant in the numerator of (2) is a linear combination of $\vec{X}^{(i)}$*-s* from the top row (see "The function is well-defined" section). It means, we can work with components of $\vec{X}^{(i)}$*-s* separately, i.e.

$$\det\begin{pmatrix} 0 & \vec{X}^{(1)} & \vec{X}^{(2)} & \dots & \vec{X}^{(n+1)} \\ x_1 & x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(n+1)} \\ x_2 & x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(n+1)} \\ \dots & \dots & \dots & \dots & \dots \\ x_n & x_n^{(1)} & x_n^{(2)} & \dots & x_n^{(n+1)} \\ 1 & 1 & 1 & \dots & 1 \end{pmatrix} = \sum_{k=1}^{n} \vec{e}_k \det\begin{pmatrix} 0 & X_k^{(1)} & X_k^{(2)} & \dots & X_k^{(n+1)} \\ x_1 & x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(n+1)} \\ x_2 & x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(n+1)} \\ \dots & \dots & \dots & \dots & \dots \\ x_n & x_n^{(1)} & x_n^{(2)} & \dots & x_n^{(n+1)} \\ 1 & 1 & 1 & \dots & 1 \end{pmatrix},$$

where $\vec{e}_k$ are coordinate vectors of the space $\vec{X}^{(i)}$-s "live in".

It's natural to try out $\vec{X}^{(i)}$-s vectors from different linear spaces — in dimensions greater or smaller than n. This is one way to extend (2), but there is another possible tweak we can make that has to do with homogeneous coordinates. Let's change vectors in the top row to $\vec{\chi}^{(i)} = \vec{X}^{(i)} + \vec{e}_{n+1}$ (add $n+1$-th dimension and set $\vec{\chi}_{n+1}^{(i)}$ to 1). Components of $\vec{\chi}^{(i)}$ are homogeneous coordinates of $\vec{X}^{(i)}$ now

$$\vec{\chi}^{(i)} = \begin{pmatrix} X_1^{(i)} \\ X_2^{(i)} \\ \dots \\ X_n^{(i)} \\ 1 \end{pmatrix}.$$

We can rewrite (2) as follows

$$\vec{\chi}(\vec{x}) = \det\begin{pmatrix} 0 & \vec{\chi}^{(1)} & \vec{\chi}^{(2)} & \dots & \vec{\chi}^{(n+1)} \\ x_1 & x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(n+1)} \\ \dots & \dots & \dots & \dots & \dots \\ x_n & x_n^{(1)} & x_n^{(2)} & \dots & x_n^{(n+1)} \\ 1 & 1 & 1 & \dots & 1 \end{pmatrix}, \qquad \left(\vec{X}(\vec{x})\right)_k = \frac{(\vec{\chi}(\vec{x}))_k}{(\vec{\chi}(\vec{x}))_{n+1}}.$$

Components of $\vec{\chi}(\vec{x})$ are homogeneous coordinates of $\vec{X}(\vec{x})$. The advantage of this notation (that could be called "switch to homogeneous coordinates in the output") is that the denominator and minus sign in (2) now reside in $(\vec{\chi}(\vec{x}))_{n+1}$

$$(\vec{\chi}(\vec{x}))_{n+1} = \det\begin{pmatrix} 0 & 1 & \dots & 1 \\ x_1 & x_1^{(1)} & \dots & x_1^{(n+1)} \\ \dots & \dots & \dots & \dots \\ x_n & x_n^{(1)} & \dots & x_n^{(n+1)} \\ 1 & 1 & \dots & 1 \end{pmatrix} = -\det\begin{pmatrix} x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(n+1)} \\ x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(n+1)} \\ \dots & \dots & \dots & \dots \\ x_n^{(1)} & x_n^{(2)} & \dots & x_n^{(n+1)} \\ 1 & 1 & \dots & 1 \end{pmatrix}.$$

Top row contains 1*-s* as by our definition $\vec{\chi}_{n+1}^{(i)} = 1$.

It's tempting to think of 1*-s* in the bottom row as the last one of x*'s* homogeneous coordinates and switch to homogeneous coordinates for all vectors lifting restriction for the last component to be necessarily equal to 1. But this move is no good — our formula need to work even if the last components of vectors become different, but it won't. The reason is that transformation will not be correctly defined and will depend on the particular choice of homogeneous coordinates of $\vec{x}^{(i)}$*-s* and $\vec{X}^{(i)}$*-s*.

Volumes of simplexes

There is a nice geometric interpretation of (2). Once again we use the Laplace expansion along the row of $\vec{X}^{(i)}$*-s*, but this time we additionally permute columns in order to place $\vec{x}$ to where the removed column was. As a result of all permutations coefficient $(-1)^{i-1}$ is multiplied by $(-1)^i$ from the Laplace expansion and -1 that was in front of the fraction from the start. This results in simple 1

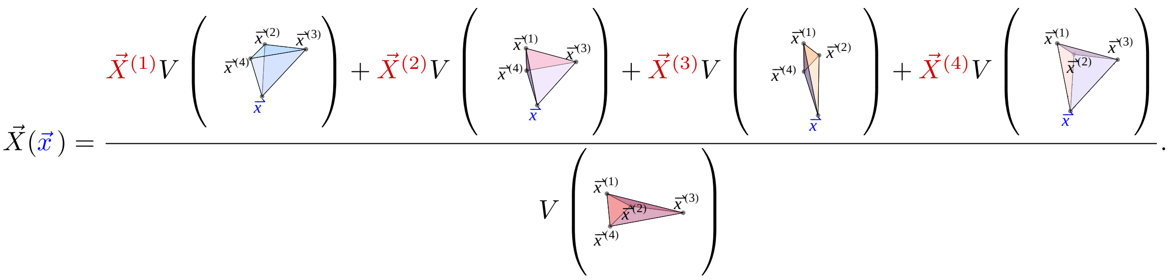
$$(-1)\det\begin{pmatrix} 0 & \vec{X}^{(1)} & \vec{X}^{(2)} & \dots & \vec{X}^{(n+1)} \\ x_1 & x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(n+1)} \\ x_2 & x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(n+1)} \\ \dots & \dots & \dots & \dots & \dots \\ x_n & x_n^{(1)} & x_n^{(2)} & \dots & x_n^{(n+1)} \\ 1 & 1 & 1 & \dots & 1 \end{pmatrix} = \vec{X}^{(1)} \underbrace{\det\begin{pmatrix} x_1 & x_1^{(2)} & \dots & x_1^{(n+1)} \\ x_2 & x_2^{(2)} & \dots & x_2^{(n+1)} \\ \dots & \dots & \dots & \dots \\ x_n & x_n^{(2)} & \dots & x_n^{(n+1)} \\ 1 & 1 & \dots & 1 \end{pmatrix}}_{V(\vec{x};\vec{x}^{\,(1)};\dots;\vec{x}^{\,(n+1)})\cdot n!} + \dots +$$

$$+\vec{X}^{(i)} \underbrace{\det\begin{pmatrix} x_1^{(1)} & \dots & x_1^{(i-1)} & x_1 & x_1^{(i+1)} & \dots & x_1^{(n+1)} \\ x_2^{(1)} & \dots & x_2^{(i-1)} & x_2 & x_2^{(i+1)} & \dots & x_2^{(n+1)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ x_n^{(1)} & \dots & x_n^{(i-1)} & x_n & x_n^{(i+1)} & \dots & x_n^{(n+1)} \\ 1 & \dots & 1 & 1 & 1 & \dots & 1 \end{pmatrix}}_{V(\vec{x}^{\,(1)};\dots;\vec{x}^{\,(i-1)};\vec{x};\vec{x}^{\,(i+1)};\dots;\vec{x}^{\,(n+1)})\cdot n!} + \dots + \vec{X}^{(n+1)} \underbrace{\det\begin{pmatrix} x_1^{(1)} & \dots & x_1^{(n)} & x_1 \\ x_2^{(1)} & \dots & x_2^{(n)} & x_2 \\ \dots & \dots & \dots & \dots \\ x_n^{(1)} & \dots & x_n^{(n)} & x_n \\ 1 & \dots & 1 & 1 \end{pmatrix}}_{V(\vec{x}^{\,(1)};\dots;\vec{x}^{\,(n)};\vec{x})\cdot n!}.$$

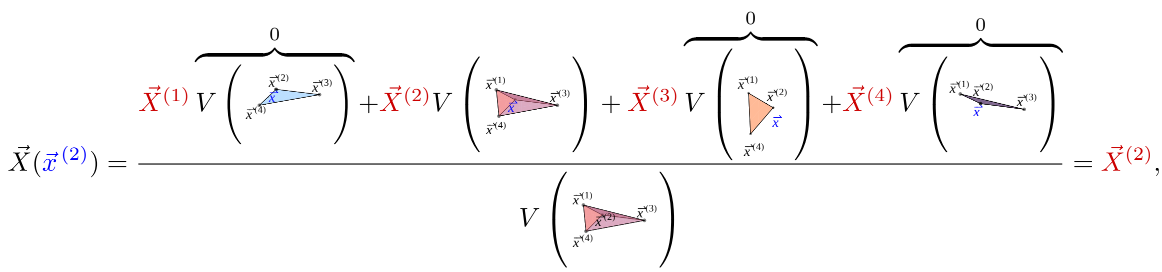
Each of the determinants in the previous equation is a directed volume of the respective parallelotope — $n!$ times the volume of the respective simplex — for brevity we designate it V, arguments in brackets appear in order corresponding vector-columns appear in the matrix under determinant. Now we see the geometrical meaning of (2) — we generate $n+1$ new simplexes from the initial simplex — the i-th new simplex is obtained by substituting $\vec{x}$ (point we want to map) instead of $\vec{x}^{\,(i)}$, then we calculate its directed volume V, divide by directed volume of the initial simplex and use this number as a weight coefficient for $\vec{X}^{(i)}$

$$\vec{X}(\vec{x}\,) = \vec{X}^{(1)} \frac{V(\vec{x};\vec{x}^{\,(2)};\dots;\vec{x}^{\,(n+1)})}{V(\vec{x}^{\,(1)};\dots;\vec{x}^{\,(n+1)})} + \cdots$$
$$\cdots + \vec{X}^{(i)} \frac{V(\vec{x}^{\,(1)};\dots;\vec{x}^{\,(i-1)};\vec{x};\vec{x}^{\,(i+1)};\dots;\vec{x}^{\,(n+1)})}{V(\vec{x}^{\,(1)};\dots;\vec{x}^{\,(n+1)})} + \cdots + \vec{X}^{(n+1)} \frac{V(\vec{x}^{\,(1)};\dots;\vec{x}^{\,(n)};\vec{x})}{V(\vec{x}^{\,(1)};\dots;\vec{x}^{\,(n+1)})}.$$

For those Readers with good spatial imagination, the last formula can be neatly visualized for 3D case



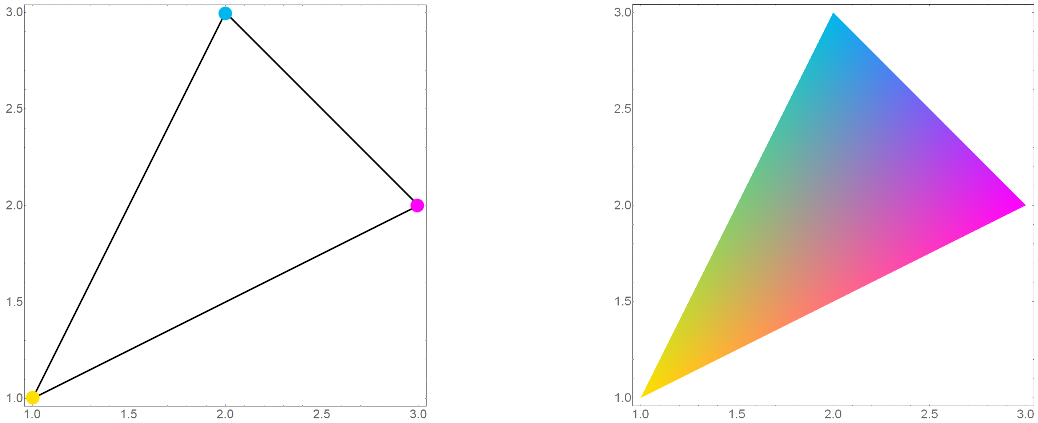
Now it's clear, why $\vec{X}$ maps $\vec{x}^{\,(i)}$ to $\vec{X}^{(i)}$ — substituting $\vec{x}^{\,(i)}$ for $\vec{x}$ "flattens out" all simplexes but the one with $\vec{x}$ in place of $\vec{x}^{\,(i)}$ (in this case we restore the initial simplex). That's why all volumes except the one are equal to zero



while the non-zero volume coincides with the volume of the initial simplex, thus is reduced by division to 1.

Multilinear interpolation

Formally, problem of multilinear interpolation can be viewed as affine transformation, but now codomain is a "space of properties" and instead of "moving points from here to there" we pull-back these properties to augment the points of the domain. A good example is multilinear interpolation of color — we can define color for vertices of certain simplex and interpolate it across the whole interior of the simplex. The result is obtained using equation (2) by setting $\vec{X}^{(i)}$-s to either numbers (grayscale) or three-component vectors (RGB). Figure 2 illustrates the principle.



(a) Initial triangle with colors assigned to vertices

(b) Color interpolated across the triangle

Figure 2: Interpolation of color across the triangle

Obviously, any other properties can be interpolated this way, but one should take care about meaningfulness of the weighted sums that appear during the process of interpolation.

Linear transformations

Obviously, linear transformations are a subset of affine transformations, but one may find an interesting inherent connection to other problems through (2) as well. First we notice, that linear transformation is affine transformation with a fixed point — origin $\vec{0}$. This fact is easy to see, since any linear transformation $\vec{L}$ should map $\vec{0}$ to $\vec{0}$ ($\vec{L}(\vec{0}) = \vec{L}(2\,\vec{0}) = 2\vec{L}(\vec{0})$), thus this is necessary. On the other hand, affine transformation (1) that maps $\vec{0}$ to $\vec{0}$ has no translation ($\vec{t} = \vec{0}$), while the leftover in form of multiplication by matrix $\hat{A}$ is linear, thus condition is sufficient.

Now we can consider (2) and take into account section "Fixed points"

$$\vec{X}(\vec{x}\,) = (-1)\frac{\det\begin{pmatrix} 0 & \vec{X}^{(1)} & \dots & \vec{X}^{(n)} & \vec{0} \\ x_1 & x_1^{(1)} & \dots & x_1^{(n)} & 0 \\ x_2 & x_2^{(1)} & \dots & x_2^{(n)} & 0 \\ \dots & \dots & \dots & \dots & \dots \\ x_n & x_n^{(1)} & \dots & x_n^{(n)} & 0 \\ 1 & 1 & \dots & 1 & 1 \end{pmatrix}}{\det\begin{pmatrix} x_1^{(1)} & \dots & x_1^{(n)} & 0 \\ x_2^{(1)} & \dots & x_2^{(n)} & 0 \\ \dots & \dots & \dots & \dots \\ x_n^{(1)} & \dots & x_n^{(n)} & 0 \\ 1 & \dots & 1 & 1 \end{pmatrix}} = (-1)\frac{\det\begin{pmatrix} 0 & \vec{X}^{(1)} & \dots & \vec{X}^{(n)} \\ x_1 & x_1^{(1)} & \dots & x_1^{(n)} \\ x_2 & x_2^{(1)} & \dots & x_2^{(n)} \\ \dots & \dots & \dots & \dots \\ x_n & x_n^{(1)} & \dots & x_n^{(n)} \end{pmatrix}}{\det\begin{pmatrix} x_1^{(1)} & \dots & x_1^{(n)} \\ x_2^{(1)} & \dots & x_2^{(n)} \\ \dots & \dots & \dots \\ x_n^{(1)} & \dots & x_n^{(n)} \end{pmatrix}}. \tag{5}$$

The latter means, we need only n non-zero points to recover linear transformation — the role of one more point goes to axes origin $\vec{0}$ (of course, all $\vec{x}^{(i)}$-s should be linearly independent).

Two matrices notation

We already know that (5) defines a unique affine transform (section "How many points do we need?") that can be rewritten in canonical form (section "Canonical notation"). But there is one more convenient form we can rewrite (5) into

$$\vec{X}(\vec{x}\,) = \underbrace{\begin{pmatrix} \boxed{\vec{X}^{(1)}} & \boxed{\vec{X}^{(2)}} & \cdots & \boxed{\vec{X}^{(n)}} \end{pmatrix}}_{\hat{X}} \underbrace{\begin{pmatrix} \boxed{\vec{x}^{\,(1)}} & \boxed{\vec{x}^{\,(2)}} & \cdots & \boxed{\vec{x}^{\,(n)}} \end{pmatrix}^{-1}}_{\hat{x}^{-1}} \begin{pmatrix} \boxed{\vec{x}} \end{pmatrix}, \tag{6}$$

where matrices are constructed from vector-columns $\vec{x}^{\,(i)}$ and $\vec{X}^{\,(i)}$ and act on vector-column $\vec{x}$. Please note, $\hat{x}$ is always square and its columns are linearly independent (otherwise (5) will not work), thus we can always find its inverse. Since matrix multiplication is always defined, we can state that linear transformation presented as (5) can always be rewritten in the mentioned form.

Now let's quickly show that (6) and (5) define the same transformation to finish the proof. First, note how matrices act on unit vectors $\vec{e}^{\,(i)}$, namely $\hat{X}\vec{e}_i = \vec{X}^{(i)}$ and $\hat{x}\vec{e}^{\,(i)} = \vec{x}^{\,(i)}$. We are interested in inverse matrix $\hat{x}^{-1}$ and we know it should do "the opposite" to what $\hat{x}$ does, namely $\hat{x}^{-1}\vec{x}^{\,(i)} = \vec{e}^{\,(i)}$. Now we can state that

$$\vec{X}(\vec{x}^{\,(i)}) = \hat{X}\hat{x}^{-1}\vec{x}^{\,(i)} = \hat{X}\vec{e}^{\,(i)} = \vec{X}^{(i)}.$$

From sections "How many points do we need?" and "Linear transformations" we know that such transformation should be unique, thus both (5) and (6) define the same linear transformation.

One can write similar to (6) expression for affine transformation (2), but he should carefully treat ones at the bottom row. We did not elaborate on this more general form for brevity.

Change of basis

Linear transformations are often mentioned in the context of basis change. If we consider $\vec{x}^{\,(i)}$ and $\vec{X}^{(i)}$ to be the same vector in different coordinate systems, we will be able to use equation from "Linear transformations" to solve the problem of transferring vector between two coordinate systems, when certain examples of such transfer (i.e. $\vec{x}^{\,(i)}$ to $\vec{X}^{(i)}$) are known.

The most frequent example of such a problem is switching to new coordinates, when coordinates of new basis vectors in old coordinate system are known. Suppose $\vec{e}^{\,(1)}, \vec{e}^{\,(2)}, \cdots, \vec{e}^{\,(n)}$ are the new basis vectors given in the old coordinate system. In the new coordinate system their coordinates should be

$$\begin{pmatrix} 1 \\ 0 \\ \cdots \\ 0 \end{pmatrix}', \begin{pmatrix} 0 \\ 1 \\ \cdots \\ 0 \end{pmatrix}', \cdots, \begin{pmatrix} 0 \\ 0 \\ \cdots \\ 1 \end{pmatrix}',$$

as they are the new orths. Primes show that coordinates are given in the new basis.

Now we want to express vector $\vec{x}$, given in old coordinate system, in the new basis. Using (5) we can write

$$\begin{pmatrix} x'_1 \\ x'_2 \\ \cdots \\ x'_n \end{pmatrix}' = (-1)\frac{\det\begin{pmatrix} 0 & \begin{pmatrix} 1 \\ 0 \\ \cdots \\ 0 \end{pmatrix}' & \begin{pmatrix} 0 \\ 1 \\ \cdots \\ 0 \end{pmatrix}' & \cdots & \begin{pmatrix} 0 \\ 0 \\ \cdots \\ 1 \end{pmatrix}' \\ x_1 & e_1^{(1)} & e_1^{(2)} & \cdots & e_1^{(n)} \\ x_2 & e_2^{(1)} & e_2^{(2)} & \cdots & e_2^{(n)} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_n & e_n^{(1)} & e_n^{(2)} & \cdots & e_n^{(n)} \end{pmatrix}}{\det\begin{pmatrix} e_1^{(1)} & e_1^{(2)} & \cdots & e_1^{(n)} \\ e_2^{(1)} & e_2^{(2)} & \cdots & e_2^{(n)} \\ \cdots & \cdots & \cdots & \cdots \\ e_n^{(1)} & e_n^{(2)} & \cdots & e_n^{(n)} \end{pmatrix}},$$

where primes designate coordinates in the new basis. Please note that the latter is a specific case of (5).

Cramer's rule

Now consider a system of linear equations

$$\hat{A}\vec{x} = \vec{b}.$$

If the system has unique solution, it can be written as

$$\vec{x} = \hat{A}^{-1}\vec{b}.$$

Suppose matrix $\hat{A}$ has entries

$$\hat{A} = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \multicolumn{4}{c}{\dotfill} \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{pmatrix}$$

and consider its action on coordinate vectors $\vec{e}_i = (0; \dots; 0; 1; 0; \dots; 0)^T$ (1 is in i-th position)

$$\hat{A}: \begin{pmatrix} 1 \\ 0 \\ \dots \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} a_{1,1} \\ a_{2,1} \\ \dots \\ a_{n,1} \end{pmatrix}, \; \hat{A}: \begin{pmatrix} 0 \\ 1 \\ \dots \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} a_{1,2} \\ a_{2,2} \\ \dots \\ a_{n,2} \end{pmatrix}, \; \dots, \; \hat{A}: \begin{pmatrix} 0 \\ 0 \\ \dots \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} a_{1,n} \\ a_{2,n} \\ \dots \\ a_{n,n} \end{pmatrix}.$$

Now it's obvious, inverse transformation will map everything "in the opposite direction"

$$\hat{A}^{-1}: \begin{pmatrix} a_{1,1} \\ a_{2,1} \\ \dots \\ a_{n,1} \end{pmatrix} \mapsto \begin{pmatrix} 1 \\ 0 \\ \dots \\ 0 \end{pmatrix}, \; \hat{A}^{-1}: \begin{pmatrix} a_{1,2} \\ a_{2,2} \\ \dots \\ a_{n,2} \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ 1 \\ \dots \\ 0 \end{pmatrix}, \; \dots, \; \hat{A}^{-1}: \begin{pmatrix} a_{1,n} \\ a_{2,n} \\ \dots \\ a_{n,n} \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ 0 \\ \dots \\ 1 \end{pmatrix},$$

which means we now have n points with their images as needed for linear transformation recovery (see section "Linear transformations"), thus we can write solution as

$$\vec{x} = \hat{A}^{-1}\vec{b} = (-1)\frac{\det\begin{pmatrix} 0 & \vec{e}_1 & \vec{e}_2 & \dots & \vec{e}_n \\ b_1 & a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ b_2 & a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \multicolumn{5}{c}{\dotfill} \\ b_n & a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{pmatrix}}{\det\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \multicolumn{4}{c}{\dotfill} \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{pmatrix}}.$$

If we use Laplace expansion along the first row, it becomes obvious that presented equation is equivalent to Cramer's rule, i.e. the last equation encapsulates all Cramer's equations.

It is interesting to notice the similarity between obtained form of Cramer's rule and previously presented equation for barycentric coordinates (4)— Cramer's rule seems to be finding barycentric coordinates of point $\vec{b}$ with respect to simplex that has $\vec{0}$ as one of its vertices (the rest coordinates are columns of matrix $\hat{A}$) and disregarding the last barycentric coordinate λ_{n+1}.

Lagrange interpolation

Let's consider a problem of polynomial interpolation, i.e. finding coefficients of the polynomial

$$y = a_1 x^n + a_2 x^{n-1} + \dots + a_n x + a_{n+1},$$

when $n+1$ pairs of x_i and corresponding y_i are known. Lagrange interpolation can be written as

$$y(x) = \sum_{i=1}^{n+1} y_i \ell_i(x), \quad \ell_i(x) = \prod_{\substack{1 \le m \le n+1 \\ m \ne i}} \frac{x - x_m}{x_i - x_m},$$

where $\ell_j(x)$ is Lagrange polynomial. Worth noting, resulting polynomial should have the lowest possible degree, e.g. a_1 can be equal to zero if degree $n-1$ is enough to interpolate given points.

Now we switch gears for a moment to determinant of the Vandermonde matrix

$$V(x_1;\ldots;x_{i-1};x_i;x_{i+1};\ldots;x_{n+1}) = \det\begin{pmatrix} x_1^n & \ldots & x_{i-1}^n & x_i^n & x_{i+1}^n & \ldots & x_{n+1}^n \\ x_1^{n-1} & \ldots & x_{i-1}^{n-1} & x_i^{n-1} & x_{i+1}^{n-1} & \ldots & x_{n+1}^{n-1} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ x_1 & \ldots & x_{i-1} & x_i & x_{i+1} & \ldots & x_{n+1} \\ 1 & \ldots & 1 & 1 & 1 & \ldots & 1 \end{pmatrix}.$$

There are two things to mention: (1) Laplace expansion along the i-th column tells us that $V(x_1;\ldots;x_{n+1})$ is a polynomial with respect to x_i of degree n, (2) when two columns are equal the determinant is zero — this tells us that $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_{n+1}$ are all the polynomial's n roots. Since every polynomial can be unambiguously written in terms of its roots we have

$$V(x_1;\ldots;x_{i-1};x_i;x_{i+1};\ldots;x_{n+1}) = C(x_1;\ldots;x_{i-1};x_{i+1};\ldots;x_{n+1}) \prod_{\substack{1\le m\le n+1\\ m\ne i}} (x_i - x_m),$$

where C is some constant depending on all $x_{j\ne i}$. Of course, nothing changes if we write x instead of x_i. The latter means

$$\ell_i(x) = \prod_{\substack{1\le m\le n+1\\ m\ne i}} \frac{x - x_m}{x_i - x_m} = \frac{V(x_1;\ldots;x_{i-1};x;x_{i+1};\ldots;x_{n+1})}{V(x_1;\ldots;x_{i-1};x_i;x_{i+1};\ldots;x_{n+1})},$$

thus Lagrange interpolation can be written as

$$y(x) = y_1\frac{V(x;x_2;\ldots;x_{n+1})}{V(x_1;\ldots;x_{n+1})} + \cdots$$
$$\cdots + y_i\frac{V(x_1;\ldots;x_{i-1};x;x_{i+1};\ldots;x_{n+1})}{V(x_1;\ldots;x_{n+1})} + \cdots + y_{n+1}\frac{V(x_1;\ldots;x_n;x)}{V(x_1;\ldots;x_{n+1})},$$

or we can do even better by employing methods of section "Volumes of simplexes" (we had almost the same expression there) and write

$$y(x) = (-1)\frac{\det\begin{pmatrix} 0 & y_1 & y_2 & \ldots & y_{n+1} \\ x^n & x_1^n & x_2^n & \ldots & x_{n+1}^n \\ x^{n-1} & x_1^{n-1} & x_2^{n-1} & \ldots & x_{n+1}^{n-1} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ x & x_1 & x_2 & \ldots & x_{n+1} \\ 1 & 1 & 1 & \ldots & 1 \end{pmatrix}}{\det\begin{pmatrix} x_1^n & x_2^n & \ldots & x_{n+1}^n \\ x_1^{n-1} & x_2^{n-1} & \ldots & x_{n+1}^{n-1} \\ \ldots & \ldots & \ldots & \ldots \\ x_1 & x_2 & \ldots & x_{n+1} \\ 1 & 1 & \ldots & 1 \end{pmatrix}},$$

that is very similar to (2) and means Lagrange interpolation is another case of simplex mapping.

Conclusions

Starting with intuition and a handful of practical objectives we have reformulated them into the problem of finding the affine transform that maps vertices of a simplex into certain points. The problem was solved with a neat formula (2), we have derived, that reminds of cross-product expression and is easy to remember. Expansions along the first row and the first column give two different insights into the formula — "barycentric" and "affine". If one needs canonical form of affine mapping, we also provide a way to derive it from presented equation (2). In a sense we presented a way to define affine (and linear) transformations not in a form of a matrix and translation vector but as a ratio of two determinants and we have also shown how to switch back and forth between our representation of mappings and the canonical one. In some applications this representation turns out to be more convenient than canonical one.

We analyzed the formula and showed its connection to barycentric and homogeneous coordinates, multilinear and Lagrange interpolations, Cramer's rule and volume of simplexes. All the examples show that (2) allows for compact and convenient notation in many cases, e.g. Cramer's method as a single expression, barycentric coordinates without explicit matrix inversion, new form of Lagrange interpolation, and different inverse problems — everything contained within one expression.

We hope that (2) will find its place in the mathematician's toolbox.