Linear Algebra
C Term, Sections C01-C04
W. J. Martin
February 8, 2010

The New French Curve

When I was in high school, drawing houses and space shuttles, we all used this cool plastic stencil called a "french curve". Here's an example[1] borrowed from geometrix.co.uk.

Figure 1: A pre-CAD/CAM French curve.

In 1968, Pierre Bézier, an engineer for Renault, developed a CAD/CAM system called UNISURF. He needed a mechanism for constructing smooth curves that passed through given points with given tangent lines. These "Bézier curves" are now fundamental in computer graphics. It is interesting to note that Bézier was not a mathematician, but rather obtained his mechanical engineering degree in 1930. Still, he remembered enough linear algebra at the age of 58 to come up with this very beautiful invention.

In this short note, we will look only at a special case.

Let $\mathbf{a}$, $\mathbf{b}$, $\mathbf{u}$ and $\mathbf{v}$ be vectors in $\mathbb{R}^2$. We aim to find a parametric curve $C = (p_x(t), p_y(t))$ (here $0 \le t \le 1$) in the plane which satisfies the following conditions

- C passes through $\mathbf{a}$ at time $t = 0$;
- C passes through $\mathbf{b}$ at time $t = 1$;
- the tangent line to C at point $\mathbf{a}$ is parallel to the vector $\mathbf{u} - \mathbf{a}$;

[1] Actually, this one is apparently a "spanish curve", but I guess those Europeans are all pretty much curved the same way.

- the tangent line to C at point $\mathbf{b}$ is parallel to the vector $\mathbf{v} - \mathbf{b}$.

To find this curve, we perform the following steps:

- let $\mathbf{r}_1$ be a point moving at constant velocity from $\mathbf{a}$ (at time $t = 0$) to $\mathbf{u}$ (at time $t = 1$);
- let $\mathbf{r}_2$ be a point moving at constant velocity from $\mathbf{u}$ (at time $t = 0$) to $\mathbf{v}$ (at time $t = 1$);
- likewise $\mathbf{r}_3$ is to be a point moving at constant velocity from $\mathbf{v}$ (at time $t = 0$) to $\mathbf{b}$ (at time $t = 1$).

Now we introduce two more vectors $\mathbf{s}_1$ and $\mathbf{s}_2$:

- at time t, the vector $\mathbf{s}_1$ takes an intermediate position between the current $\mathbf{r}_1$ and the current $\mathbf{r}_2$:
$$\mathbf{s}_1 = (1-t)\mathbf{r}_1 + t\mathbf{r}_2 \ ;$$
- at time t, the vector $\mathbf{s}_2$ takes an analogous intermediate position between the current $\mathbf{r}_2$ and the current $\mathbf{r}_3$:
$$\mathbf{s}_2 = (1-t)\mathbf{r}_2 + t\mathbf{r}_3 \ ;$$

Finally, we get the vector we are interested in:

- the point $\mathbf{p}$ on the actual curve at time t is given by
$$\mathbf{p} = (1-t)\mathbf{s}_1 + t\mathbf{s}_2.$$

Here's how to represent these time-dependent vectors algebraically. At time t, we have $\mathbf{r}_1$ expressed as a linear combination: $\mathbf{r}_1 = (1-t)\mathbf{a} + t\mathbf{u}$. The same sort of relationship will hold for the remaining five vectors. We have

$$\begin{aligned} \mathbf{r}_1 &= (1-t)\mathbf{a} + t\mathbf{u} \\ \mathbf{r}_2 &= (1-t)\mathbf{u} + t\mathbf{v} \\ \mathbf{r}_3 &= (1-t)\mathbf{v} + t\mathbf{b} \end{aligned}$$

The next vectors are not moving at constant velocity, but have equally simple algebraic expressions:

$$\begin{aligned} \mathbf{s}_1 &= (1-t)\mathbf{r}_1 + t\mathbf{r}_2 \\ \mathbf{s}_2 &= (1-t)\mathbf{r}_2 + t\mathbf{r}_3 \\ \mathbf{p} &= (1-t)\mathbf{s}_1 + t\mathbf{s}_2 \end{aligned}$$

From the above six equations, we use substitution and simplification to express the point $\mathbf{p}$ as a linear combination of the original four vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{u}$ and $\mathbf{v}$. Each coefficient will be some function of the parameter t.

$$\begin{aligned}
\mathbf{p} &= (1-t)\mathbf{s}_1 + t\mathbf{s}_2 \\
&= (1-t)[(1-t)\mathbf{r}_1 + t\mathbf{r}_2] + t[(1-t)\mathbf{r}_2 + t\mathbf{r}_3] \\
&= (1-t)^2\mathbf{r}_1 + 2t(1-t)\mathbf{r}_2 + t^2\mathbf{r}_3 \\
&= (1-t)^2[(1-t)\mathbf{a} + t\mathbf{u}] + 2t(1-t)[(1-t)\mathbf{u} + t\mathbf{v}] + t^2[(1-t)\mathbf{v} + t\mathbf{b}] \\
&= (1-t)^3\mathbf{a} + 3t(1-t)^2\mathbf{u} + 3t^2(1-t)\mathbf{v} + t^3\mathbf{b}
\end{aligned}$$

So we obtain Bézier's magic formula for the parametric curve:

$$\mathbf{p}(t) = (1-t)^3\mathbf{a} + 3t(1-t)^2\mathbf{u} + 3t^2(1-t)\mathbf{v} + t^3\mathbf{b}.$$

At any point in time t, the vector $\mathbf{p}$ is a special linear combination of the four control points. This generalizes naturally to more complicated Bézier curves where any number of control points can be provided by the user.

EXAMPLE 1: Suppose, in the above setup, that $\mathbf{a} = (0,0)$, $\mathbf{b} = (2,4)$, $\mathbf{u} = (\frac{2}{3},0)$ and $\mathbf{v} = (\frac{4}{3},-4)$. Let's use the result above to find a parametric form for the Bézier curve.

SOLUTION: We are given specific points:

$$\mathbf{a} = (0,0), \qquad \mathbf{u} = (\frac{2}{3},0), \qquad \mathbf{v} = (\frac{4}{3},-4), \qquad \mathbf{b} = (2,4).$$

We substitute these values and find parametric equations

$$\begin{aligned}
\mathbf{p} &= (1-t)^3\mathbf{a} + 3t(1-t)^2\mathbf{u} + 3t^2(1-t)\mathbf{v} + t^3\mathbf{b} \\
&= (1-t)^3\begin{bmatrix} 0 \\ 0 \end{bmatrix} + 3t(1-t)^2\begin{bmatrix} 2/3 \\ 0 \end{bmatrix} + 3t^2(1-t)\begin{bmatrix} 4/3 \\ -4 \end{bmatrix} + t^3\begin{bmatrix} 2 \\ 4 \end{bmatrix} \\
&= \begin{bmatrix} 2t(1-t)^2 + 4t^2(1-t) + 2t^3 \\ -12t^2(1-t) + 4t^3 \end{bmatrix} \\
&= \begin{bmatrix} 2t \\ -12t^2 + 16t^3 \end{bmatrix}
\end{aligned}$$

So, with the substitution $t = x/2$, we find that the curve has equation

$$y = 2x^3 - 3x^2.$$

Note that, we cannot always write the curve in this form. It just so happens that this example has y as a function of x.

EXAMPLE 2: Now suppose you start with control points

$$\mathbf{a} = (0, 6), \qquad \mathbf{u} = (6, 4), \qquad \mathbf{v} = (0, 2), \qquad \mathbf{b} = (6, 0).$$

SOLUTION: Then the Bézier curve is

$$\begin{aligned} \mathbf{p}(t) &= (1-t)^3\mathbf{a} + 3t(1-t)^2\mathbf{u} + 3t^2(1-t)\mathbf{v} + t^3\mathbf{b} \\ p_x(t) &= 0(1-t)^3 + 3 \cdot 6 \cdot t(1-t)^2 + 3 \cdot 0 \cdot t^2(1-t) + 6t^3 \\ &= 18t - 36t^2 + 24t^3 \\ p_y(t) &= 6(1-t)^3 + 3 \cdot 4 \cdot t(1-t)^2 + 3 \cdot 2 \cdot t^2(1-t) + 0t^3 \\ &= 6 - 6t \end{aligned}$$

so that, in this case, x can be expressed as a polynomial function of y. You can guess that this is again very special; if you move $\mathbf{b}$ to $(4, 6)$, for example, neither x nor y can be expressed as a function of the other.

EXAMPLE 3: Now suppose you start with points $\mathbf{a} = (-2, 4)$ and $\mathbf{b} = (3, 9)$ and you want the Bézier curve to be exactly the curve $y = x^2$. Can we use linear algebra to determine the correct control points $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$ in this case?