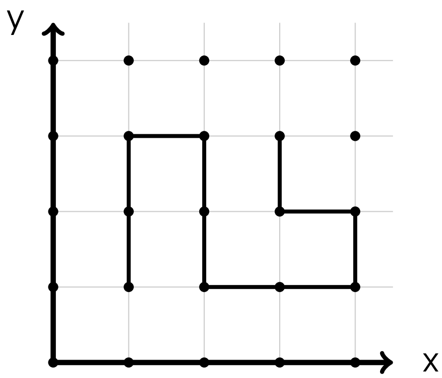
Curve Fitting

John D Mangual

I feel there are ways to do curve-fitting that look up and down the ladder of abstractions. Let's start with something pretty vanilla:



Can we draw a curve that passes through all these points? There are two strategies to try to hook around this many constraints:

- Lagrange Interpolation
- Bezier Curves

And maybe it depends on the type of constraint problem you are trying to solve:

- passing through points
- weaving arround points

I think one way to motivate a theory is to start with a problem you want to solve.[1] We like to brag about how good we are at weaving around constraints. How bad can we be? Let $f : [0, 1] \to \mathbb{R}$ be a real-valued function:

$$f(x) = \begin{cases} 1 & x \notin \mathbb{Q} \\ 0 & x \in \mathbb{Q} \end{cases}$$

Then we can ask what the integral is between 0 and 1. Overwhelmingly, the answer should be:

$$\int_0^1 f(x)\, dx = 1 \times \Big| \{x \notin \mathbb{Q}\} \cap [0, 1] \Big| + 0 \times \Big| \{x \in \mathbb{Q}\} \cap [0, 1] \Big| = 1$$

Relatively innocent-functions like these are sufficient Riemann integration. I reasoned there are vastly more irrational numbers than rational numbers.

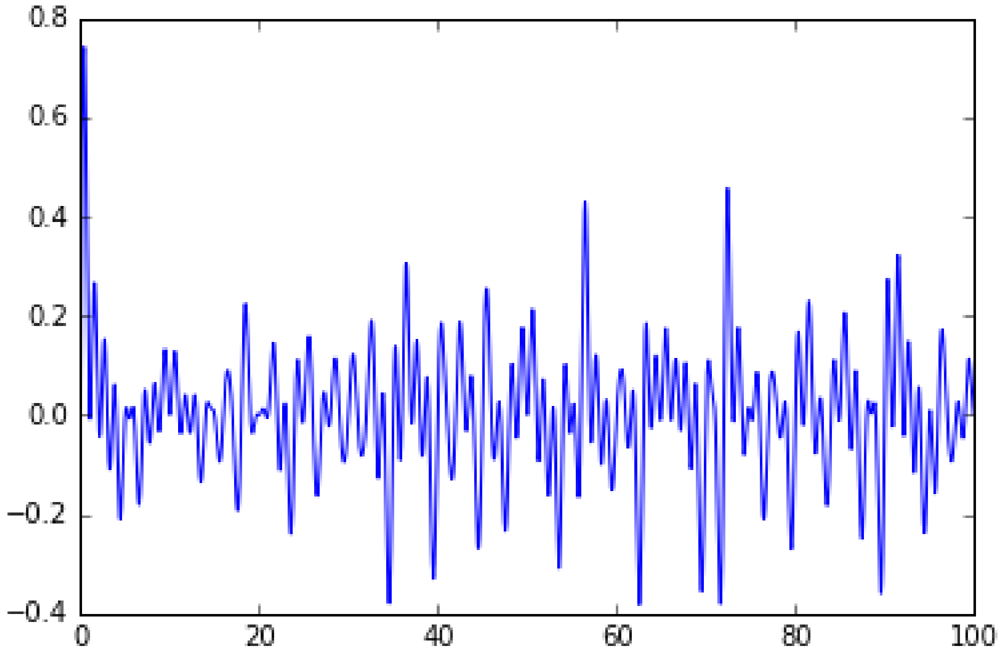
[1]Other times, theories comes out of the box, complete. I think such theories are unusuable, or they leave from for input from you and me. A comeback from that camp could be, "John, why are you so obsessed with this particular problem?" And I don't have any good reason. Just because. Sometimes about me, and the time and place makes me interested. And I could be wrong!

This turns out to be one of the worse functions around, because it looks like $\mathbf{1}$ but isn't:

$$f(x) \approx \mathbf{1} \quad \text{therefore} \quad \int f \approx \int \mathbf{1}$$

If I recall, we placed intervals around every single fraction $\frac{a}{b} \in \mathbb{Q}$ we get an upper estimate for how large this integral could be, and that upper estimate $\to 0$.

How to deal with a function that often looks like $f(x) \equiv 1$ but isn't?[2]



Here is a plot if we add all of the pure tones at all frequencies $\frac{a}{b}$ with $0 < a < b < 10$. Theoretically we have found:

$$\frac{1}{28} \sum_{0<a<b<10} e^{2\pi i \frac{a}{b} t} \approx \frac{1}{c\, N^2} \sum_{0<a<b<N} e^{2\pi i \frac{a}{b} t} = \begin{cases} 1 & \text{if } t = 0 \\ 0 & \text{if } t \neq 0 \end{cases}$$

The constant in front is an oversimplification. The odds of two numbers being relatively primes is a famous one:

$$\phi(1) + \phi(2) + \cdots + \phi(N) = \#\{(a,b) : a < b \text{ and } \gcd(a,b) = 1\} \approx \frac{2}{\zeta(2)^2}$$

Analysis is the branch of math where we account for all the uses of the $\approx$ symbol. What if I told you the function we just charted is $\approx \mathbf{0}$?

[2]One set of problems that plagued me was if we had two competing definition of limit, maybe one returns a number and the other does not. If we have two limiting procedures 1 and 2, maybe:

$$\big[\lim\big]_1 a_n = A \quad \text{implies} \quad \big[\lim\big]_2 a_n = A$$

Most of the time we don't really care how the limit procedure is defined. Who cares really?

$$N \gg 1 \quad \text{implies} \quad a_N \approx A$$

and the impliciation looks self-evident and most people won't question it. The only reason we remember the exception is because that particular conversation went on record.

...

References

(1) ...