Convolution

Recall the definition of *cross-correlation* from lecture 4 which I write here for 1D functions:

$$f(x) \otimes I(x) \equiv \sum_u f(u-x)\ I(u)\ .$$

Convolution is defined slightly differently, namely:

$$f(x) * I(x) \ \equiv \ \sum_u f(x-u)\ I(u).$$

Note that the only difference here is that the argument of $f(\ )$ is now flipped. So whenever we have a cross-correlation, we can think of it as a convolution with a flipped function, and vice-versa. Also note that if $f()$ happens to be symmetric (like a Gaussian), then there is no difference between convolution and cross-correlation.

In general though, there is a difference in how we think of cross correlation and convolution. We think of cross-correlation as sliding a template function f across another function and taking the inner product. We think of convolution $f(x) * I(x)$ as adding up shifted versions of the function $f(x)$, namely $f(x-u)$. Each shifted version is weighted by the value $I(u)$, where u is the shift.

ASIDE: The question comes up of what to do when $f(x-u)$ is not defined for some value of $x-u$. This should be familiar to you, since a similar problem arose when we defined cross-correlation in lecture 4, and it has come up in assignments. Here we can do the same thing as we did there, and just 'zero-pad' the function $I()$ beyond the domain where it is defined. An alternative, which we will mention later is to treat $I()$ as periodic.

Algebraic properties of convolution

One surprising and useful property the convolution operation is that it *commutative*: one can switch the order of the two functions I and f in the convolution without affecting the result. The property does *not* always hold for cross-correlation.

To prove that convolution is commutative, we pad $I(x)$ and $f(x)$ with zeros. This allows us to take the summation from $-\infty$ to ∞.

$$I(x) * f(x) = \sum_{u=-\infty}^{\infty} f(x-u)I(u)$$

Using the substitution $w = x - u$, we

$$I(x) * f(x) = \sum_{w=-\infty}^{\infty} f(w)I(x-w) \ = I(x) * f(x)$$

If you think of I as a signal and f as a filter then you don't need to be concerned about order of writing $I * f$ or $f * I$ since they are the same.

A second important property of convolution is that it is *associative*:

$$I \ * \ (f_1 \ * \ f_2) \ = \ (I \ * \ f_1) \ * \ f_2$$

Again the proof is simple, and you should work it out for yourself.

Why are these properties useful? Often, in signal processing, we perform a sequence of operations. For example, you might average the pixels in a local neighborhood, then take their derivative (or second derivative). The algebraic properties just described give us some flexibility in the order of operations. For example, suppose we blur an image $I(x)$ and then take its local difference. We get the equivalent result if we take the local difference on the blur function and convolve the result with the image:

$$(D(x) * B(x) \ ) \ * I(x) = D(x) \ * (B(x) \ * \ I(x)).$$

One final property of convolution is that it is *distributive*:

$$(I_1 + I_2) \ * \ f \ = \ I_1 * f \ + \ I_2 * f$$

This is also simple to prove and I leave it to you as an exercise. This property is also useful. For example, if $I_1 = I(x)$ is an image and $I_2 = n(x)$ is a noise function added to the image, then if we blur the "image+noise," we get the same result as if we blur the image and noise separately, and then add the results together.

Impulse functions, and impulse response function

Define a "delta" function

$$\delta(x) = \begin{cases} 1, & x = 0 \\ 0, & \text{otherwise.} \end{cases}$$

$\delta(x)$ is also known as an *impulse function.* It is straightforward to show that, for any function $I(x)$,

$$I(x) = \delta(x) * I(x).$$

Another way to interpret the above equation is to think of a function $I(x)$ as a sum of delta functions

$$I(x) = \sum_{u=-\infty}^{\infty} \delta(x-u) I(u) \ ,$$

namely, if we put a delta function at each value of u and multiply that delta function by the value $I(u)$, then we get the original function.

Finally, suppose we have a mapping ("convolve with $f(x)$")

$$I(x) \rightarrow I(x) * f(x)$$

In this case, we often refer to $f(x)$ as an *impulse response function.* The reason is that if $I(x)$ were an impulse $\delta(x)$ then it would map to $f(x)$. That is, $f(x)$ is the response (output) when the stimulus (input) is $\delta(x)$.

Recall that convolution $f(x) * I(x)$ is defined by adding up shifted versions of the function $f(x)$, where each shifted version is weighted by a value $I(u)$ where u is the shift. Thus, thinking $I(x)$ as a sum of delta functions, we see now that $f(x) * I(x)$ can be interpreted a sum of impulse response functions $f(x-u)$ shifted by different amounts u and weighted by different amounts $I(u)$.

Sinusoids and convolution

We next show that sinusoids have a special behavior under convolution. Take a cosine function with k cycles from $x = 0$ to $x = N$, where k is an integer,

$$\cos(\frac{2\pi k}{N}\ x).$$

Note that this cosine function has the same value at $x = N$ as at $x = 0$. Suppose we were to convolve the cosine with a function $h(x)$ which is defined on $x \in 0, \ldots, N-1$:

$$h(x) * \cos(k\ \frac{2\pi}{N}\ x) \ = \ \sum_{x'=0}^{N-1} h(x')\ \cos(k\frac{2\pi}{N}(x - x'))$$

[**BEGIN ASIDE** (I did not include this in the lecture slides since it is just a calculation.)]

Recalling the trigonometry identity from Calculus 1,

$$\cos(\alpha + \beta) \ = \ \cos\alpha\cos\beta - \sin\alpha\sin\beta$$

we can expand the cos() in the summation on the right side of the above equation, and so the right hand side is just a sum of sine and cosine functions with variable x and constant frequency k. Thus, it can be written

$$h(x) * \cos(k\ \frac{2\pi}{N}\ x) \ = \ a\cos(\frac{2\pi}{N}\ kx) - b\sin(\frac{2\pi}{N}\ kx) \tag{1}$$

which was **Claim 1** in the lecture slides. The values of a and b depend on k and on the function $h(x)$ as follows:

$$a = \sum_{x'=0}^{N-1} h(x')\ \cos(k\frac{2\pi}{N}x')$$

$$b = \sum_{x'=0}^{N-1} h(x')\ \sin(k\frac{2\pi}{N}x')$$

which are just the inner products of the N dimensional vectors $h(\cdot)$ with a cosine or sine of frequency k, respectively. [**END ASIDE**]

Let's simplify Eq. (1). Let (a, b) be a 2D vector, and define angle ϕ such that

$$(\cos\phi, \sin\phi) = \frac{1}{\sqrt{a^2 + b^2}}(a, b).$$

Then

$$\begin{aligned} h * \cos(k\ \frac{2\pi}{N}\ x) &= \sqrt{a^2+b^2}\ (\cos(\phi)\cos(\frac{2\pi}{N}\ kx) + \sin(\phi)\sin(\frac{2\pi}{N}\ kx)) \\ &= \sqrt{a^2+b^2}\ \cos((\frac{2\pi}{N}kx) - \phi) \end{aligned}$$

which was **Claim 2** in the lecture slides. The quantity $\sqrt{a^2 + b^2}$ is called the *amplitude* and ϕ is called the *phase*. The amplitude and phase depend on frequency k and on the function $h(\cdot)$.

To briefly summarize, we have shown that convolving a cosine with an arbitrary function $h(x)$ gives you back a cosine of the same frequency k, but with possibly different amplitude and possibly phase shifted in position x. (Exactly the same argument can be made for a sine function.) These amplitude and phase changes turn out to be very important, as we'll see in the next few weeks when we discuss sound processing by the ear.

One final point: I made **Claim 3** in the slides, namely that any function $I(x)$ can be written as a sum of sine and cosine functions:

$$I(x) = \sum_{k=0}^{\frac{N}{2}} a_k \ \cos(\frac{2\pi}{N}kx) + \sum_{k=1}^{\frac{N}{2}-1} b_k \ \sin(\frac{2\pi}{N}kx)$$

Because of time constraints and because we will not use this representation, I won't prove that claim. Instead, what I will do (next lecture) is give you an alternative representation, called the *Fourier* representation, which is slightly different. The Fourier transform requires that we use complex numbers, so I will spend the rest of the lecture reviewing the basics.

Complex numbers (review)

To decompose functions into sines and cosines we are going to use complex variables. Recall that a complex number c consists of a pair of numbers, (a, b) called the "real" and the "imaginary" part. One often writes this pair using the notation

$$c = a + bi.$$

We define *addition* of two complex numbers by adding their real and imaginary parts separately:

$$c_1 + c_2 = (a_1 + a_2, b_1 + b_2)$$

or

$$c_1 + c_2 = (a_1 + a_2) + (b_1 + b_2)i.$$

We can define *multiplication* of two complex numbers by writing the two numbers in polar coordinates:

$$c_1 = r_1(\cos\theta_1 + i \ \sin\theta_1)$$
$$c_2 = r_2(\cos\theta_2 + i \ \sin\theta_2)$$

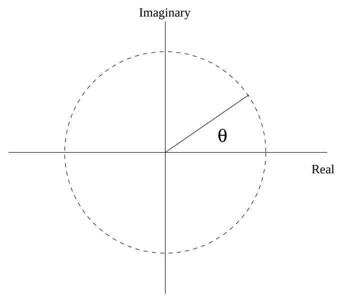
and defining the product $c_1 \ c_2$ to have a length $r_1 r_2$ and an angle $\theta_1 + \theta_2$:

$$c_1 \ c_2 = r_1 r_2(\cos(\theta_1 + \theta_2) + i \ \sin(\theta_1 + \theta_2)).$$

For example, take the case that $a = 0, b = 1$, or i. For this number, $r = 1, \theta = \frac{\pi}{2}$. So $c^2 = i^2$ has $r = 1$ and $\theta = \frac{\pi}{2} + \frac{\pi}{2} = \pi$ and so $i^2 = \cos(\pi) = -1$. Thus

$$i^2 = -1.$$

There is really nothing mysterious about this number i, once you understand that we are *defining* multiplication on pairs (a, b) of numbers in this special way.



Euler's equation

To multiply complex numbers, we often express the numbers using *Euler's equation*:

$$e^{i\theta} = \cos\theta \ + \ i \ \sin\theta$$

which represents a point on the unit circle in the complex plane.
Here are some examples:

$$e^{i0} = 1, \quad e^{i\pi/2} = i, \quad e^{i\pi} = -1, \quad e^{i\pi/4} = \frac{1}{\sqrt{2}}(1+i), \quad e^{i2\pi n} = 1 \ \text{ for any integer } n$$

More generally, consider what happens when we multiply two complex numbers $e^{i\theta_1}$ and $e^{i\theta_2}$. The definition of multiplication gives:

$$e^{i\theta_1}e^{i\theta_2} = e^{i(\theta_1+\theta_2)}$$

Using Euler's equation for the two terms on the left side gives:

$$(\cos\theta_1 + i \ \sin\theta_1)(\ \cos\theta_2 + i \ \sin\theta_2) \ = \ (\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2) + i \ (\cos\theta_1\sin\theta_2 + \sin\theta_1\cos\theta_2)$$

Using Euler's equation for the right side gives:

$$\cos(\theta_1+\theta_2) \ + \ i \ \sin(\theta_1+\theta_2).$$

Thus,

$$\cos(\theta_1+\theta_2) = \cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2$$
$$\sin(\theta_1+\theta_2) = \cos\theta_1\sin\theta_2 + \sin\theta_1\cos\theta_2$$

which are familiar trig identies that you learned in Calculus.

Complex conjugate and inverse

The *complex conjugate* of $c = a + bi$ is defined

$$\bar{c} = a - bi.$$

The complex conjugate has the property that

$$c \ \bar{c} \ = \ | \ c \ |^2 \ = \ a^2 + b^2 \ .$$

In particular, $e^{-i\theta}$ is the *complex conjugate* of $e^{i\theta}$ and

$$e^{i\theta} \, e^{-i\theta} \; = \; 1 \; .$$

The complex conjugate of c should not be confused with the inverse of c, namely the complex number c^{-1} which satisfies $cc^{-1} = 1$,

$$c^{-1} = \frac{1}{|c|}\overline{c} \; .$$

