Chapter 9

Correlation and Matched Filtering

[Updated March 8, 2014 by Aaron Lanterman.]

Suppose we measured a waveform that had the form of a signal corrupted by additive noise

$$x(t) = s_k(t) + n(t), \tag{9.1}$$

where s_1, s_2, etc. represent different kinds of signals, which we call *templates*, that we are trying to discriminate between. We are not going to say much about the noise; properly treating it requires a thorough discussion of probability and random processes, which is beyond the scope of this text.

9.1 Squared-error metrics and correlation processing

One reasonable approach might be to measure the "error" between each template and the actual measured data, and pick the template that yields the lowest error. The *squared error* is commonly employed:

$$\begin{aligned}
\int_{-\infty}^{\infty} |x(t) - s_k(t)|^2 dt &= \int_{-\infty}^{\infty} [x(t) - s_k(t)][x(t) - s_k(t)]^* dt & (9.2)\\
&= \int_{-\infty}^{\infty} [x(t) - s_k(t)][x^*(t) - s_k^*(t)] dt & (9.3)\\
&= \int_{-\infty}^{\infty} x(t)x^*(t) - s_k(t)x^*(t) - s_k^*(t)x(t) - s_k(t)s_k^*(t) dt & (9.4)\\
&= \int_{-\infty}^{\infty} |x(t)|^2 - s_k(t)x^*(t) - s_k^*(t)x(t) - |s_k(t)|^2 dt. & (9.5)
\end{aligned}$$

The squared error is mathematically convenient. It is particularly appropriate if the noise $n(t)$ is Gaussian, but you do not need to know anything about Gaussian probability distributions – or probability in general – to understand what follows.

For convenience, suppose the signal templates are all normalized to have the same energy, i.e. $\int_{-\infty}^{\infty} |s_k(t)|^2 dt$ is the same for every k. Also, note that the $\int_{-\infty}^{\infty} |x(t)|^2$ term does not depend upon k. So in trying to minimize the error with respect to k, we can drop the first and last terms in (9.5). The middle two terms consist

Figure 9.1: Correlation classifier.

of something added to its complex conjugate, so our minimization problem reduces to finding the k that minimizes

$$2\Re e\left\{-\int_{-\infty}^{\infty} x(t)s_k^*(t)\right\}, \tag{9.6}$$

or equivalently, finding the k that maximizes

$$\Re e\left\{\int_{-\infty}^{\infty} x(t)s_k^*(t)\right\} dt. \tag{9.7}$$

We dropped the 2 in front since it does not change the result of the maximization. This procedure is referred to as a *correlation classification*; it may be the most common form of basic "template matching" used in pattern recognition. Essentially, we want to take the inner product[1] of the data with each template, and find the template that produces the "best" match.[2]

9.2 Matched filter implementation

The correlator classifier computation described in (9.7) has the form of a multiplier, followed by an integrator (Figure 9.1). There is an alternative way of implementing this operation that consists of running the data through an LTI filter, and then sampling the output of the filter at a particular time. Let us reconsider that inner product operation:

$$\int_{-\infty}^{\infty} x(t)s_k^*(t)dt. \tag{9.8}$$

Define an LTI filter with the impulse response $h_k(t) = s_k^*(-t)$. Notice that $h_k(t)$ is a conjugated and time-reversed version of the template; this is the impulse response of the "matched filter." If we feed this filter the data, the output is

$$y(t) = \int_{-\infty}^{\infty} x(\tau)h_k(t-\tau)d\tau. \tag{9.9}$$

Notice that the output of the filter at time 0 is equivalent to the inner product of (9.8):

$$y(0) = \int_{-\infty}^{\infty} x(\tau)h_k(0-\tau)d\tau = \int_{-\infty}^{\infty} x(\tau)h_k(-\tau)d\tau = \int_{-\infty}^{\infty} x(\tau)s_k^*(\tau)d\tau. \tag{9.10}$$

9.3 Delay estimation

This matched filter viewpoint of Section 9.2 is particularly handy if we are looking for a time-shifted version of the template and we don't know the amount of the shift. Quite often, we may have a single "template," and the amount of the time shift (usually a delay) is what we are interested in learning. In many radar and

[1] A more mathematically thorough treatment would demonstrate that this inner product is an aspect of the "projection" of the data onto the template.

[2] We put "best" in quotes since other kinds of error functions could be used that might have a different idea of what "best" means.

sonar applications, we transmit a pulse and wait for it to bounce off of an object and return. The time delay is proportional to the range to the object. The radar and sonar signals are often "bandpass" signals with a high-frequency carrier, and hence well represented by complex baseband representations. This is the main reason we made sure the exposition in Section 9.1 worked for complex signals, and not just real signals.

Instead of considering a set of templates $s_1(t)$, $s_2(t)$, etc., suppose the data is described by $x(t) = s(t-\Delta) + n(t)$, where Δ is some unknown delay. We can imagine running an infinite set of correlation detectors, with each detector corresponding to a particular Δ. Fortunately, we do not actually need to build an infinite number of multipliers and an infinite number of integrators. If we neglect noise, the output of the matched filter in this scenario is

$$\begin{aligned} y(t) &= \int_{-\infty}^{\infty} x(\tau)h(t-\tau)d\tau = \int_{-\infty}^{\infty} s(\tau-\Delta)h(t-\tau)d\tau && (9.11) \\ &= \int_{-\infty}^{\infty} s(\tau-\Delta)s^*(-(t-\tau))d\tau = \int_{-\infty}^{\infty} s(\tau-\Delta)s^*(\tau-t)d\tau. && (9.12) \end{aligned}$$

This is a good time to introduce the Schwarz inequality, which, in this context, says

$$\left| \int_{-\infty}^{\infty} f(t)g^*(t)dt \right|^2 \leq \left[\int_{-\infty}^{\infty} |f(t)|^2 dt \right] \left[\int_{-\infty}^{\infty} |g(t)|^2 dt \right] \tag{9.13}$$

with equality if and only if $f(t) = \alpha g(t)$, for some α. If we correspond $f(t)$ with $s(\tau-\Delta)$ and $g(t)$ with $s(\tau-t)$, we see that

$$\left| \int_{-\infty}^{\infty} s(\tau-\Delta)s^*(\tau-t)d\tau \right|^2 \tag{9.14}$$

is going to be largest when $s(\tau-\Delta) = s^*(\tau-t)$, i.e. when $t = \Delta$. This gives us a procedure for finding the time-shift: filter the data with a conjugated, reverse copy of transmitted waveform, and find the t for which the energy at the output power of the filter is the largest.

The output of the matched filter for $\Delta = 0$ has a special name; it is called the *autocorrelation* of the signal:

$$\int_{-\infty}^{\infty} s(\tau)s^*(\tau-t)d\tau. \tag{9.15}$$

If we substitute $\tilde{\tau} = \tau - t$, giving $\tau = \tilde{\tau} + t$, we can rewrite the autocorrelation of $s(t)$ as

$$\int_{-\infty}^{\infty} s(\tilde{\tau}+t)s^*(\tilde{\tau})d\tilde{\tau}, \tag{9.16}$$

which is how you will see it defined in most textbooks. The autocorrelation of a waveform is basically that waveform convolved with a time-reversed version of itself. We can also define a more general cross-correlation between two functions:

$$\int_{-\infty}^{\infty} x(\tilde{\tau}+t)s^*(\tilde{\tau})d\tilde{\tau}, \tag{9.17}$$

In a simple radar ranging system, if the data contains a single "target," the output of the matched filter (not including noise) is the autocorrelation function of the transmitted waveform, time-shifted to where the target is located in time.

The autocorrelation is an important tool in waveform design for range estimation, since it characterizes many aspects of system performance. The curvature of the peak of the autocorrelation is related to how

Figure 9.2: Example of a linear FM sweep.

Figure 9.3: Autocorrelation function of the linear FM sweep in Figure 9.2.

accurately we can estimate the time shift, and hence range, under noisy conditions. If multiple targets are present, targets that are close in range may blur together and look like a single target. The broadness of the peak of the autocorrelation is indicative of the *resolution* of the system, which addresses how well we can discriminate between close targets. The sidelobes of the autocorrelation function provide a sense of how a target with weak reflectivity may "hide" in the sidelobe of a target with stronger reflectivity. This brief paragraph is intended to only provide a small taste of such issues; thorough definitions of these properties and related explorations are best found in dedicated texts on remote sensing.

Matched filtering also allows radar designers to pull off a slick trick called *pulse compression*. To obtain good range resolution, we might intuitively want our waveforms to be short in duration. But transmitters are usually limited in terms of output power, which leads us to want to use long waveforms to put lots of energy on the target to be able to combat noise. If we can find a waveform with a lengthy time extent, but whose autocorrelation has a narrow mainlobe, match filtering essentially compresses the energy over that long time extent into a narrow pulse in the matched filter output, so we can enjoy both good resolution and good signal-to-noise ratio. For instance, a linear FM waveform (Figure 9.2) has an autocorrelation function that looks somewhat sinc-like (Figure 9.3), although it is by no means a "pure" sinc.

9.4 Causal concerns

In most applications, $s(t)$ will be non-zero for some time range $0 \leq t \leq L$, for some length L, and be zero outside of that range. This means that the matched filter $h(t) = s(-t)$, as defined in the sections above, will be non-causal and hence be unable to be implemented in real-time systems. In practice, this is not cause for much concern. We can implement a *causal* matched filter $h_c(t) = s^*(-(t-L)) = s^*(L-t)$, in which we simply shift our original matched filter $h(t)$ to the right by enough time that the resulting $h_c(t)$ is causal. All of our previous results apply, except that the outputs are now delayed by L.

9.5 A caveat

The "derivations" of correlation classifiers and matched filter structures given above were somewhat heuristic, since they were intended to be digestible by readers lacking experience with probability theory and random processes. These structures are usually derived via a more detailed analysis that includes rigorously modeling the noise, in which one can make precise statements about the signal-to-noise ratio.