Matched Filtering

Material not in textbook

Dr D. Laurenson

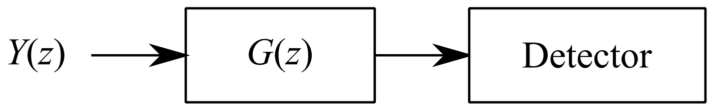
18th November 2020

1 Introduction

This material is not covered in the course textbook, so students should use these notes, and refer to on-line material should further clarification be required.

In certain problems we need to identify the presence of a known signal. In general, this is a non-trivial problem. We will first explore the use of an inverse filter approach, and highlight the issues associated with this. Then the construction of a matched filter will be derived.

We will consider a detection system where the presence of a known signal, $x(n)$, is to be identified in a signal $y(n)$. The task is to identify a suitable $G(z)$.



Applications of matched filters include:

- Communications (detecting symbols)
- System characterisation (e.g. radar or sonar)
- Biomedical signal detection (e.g. EEG)

2 Inverse Filtering

An inverse filter is defined as follows:

If $Y(z) = z^{-D}X(z) + N(z)$, then an inverse filter of $X(z)$ is:

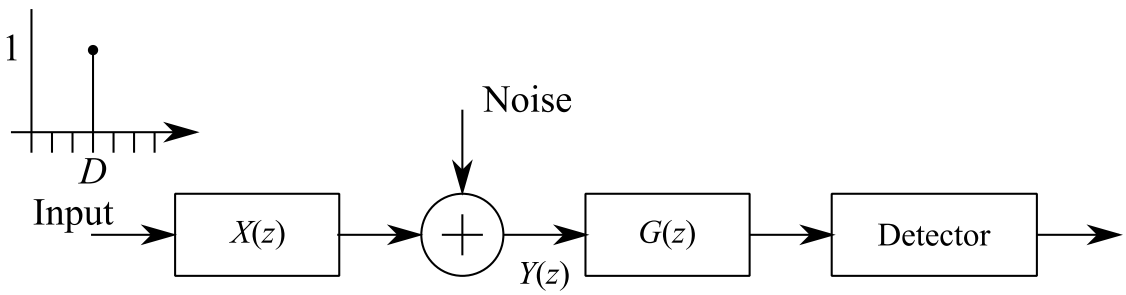
$$G(z) = \frac{1}{X(z)}$$

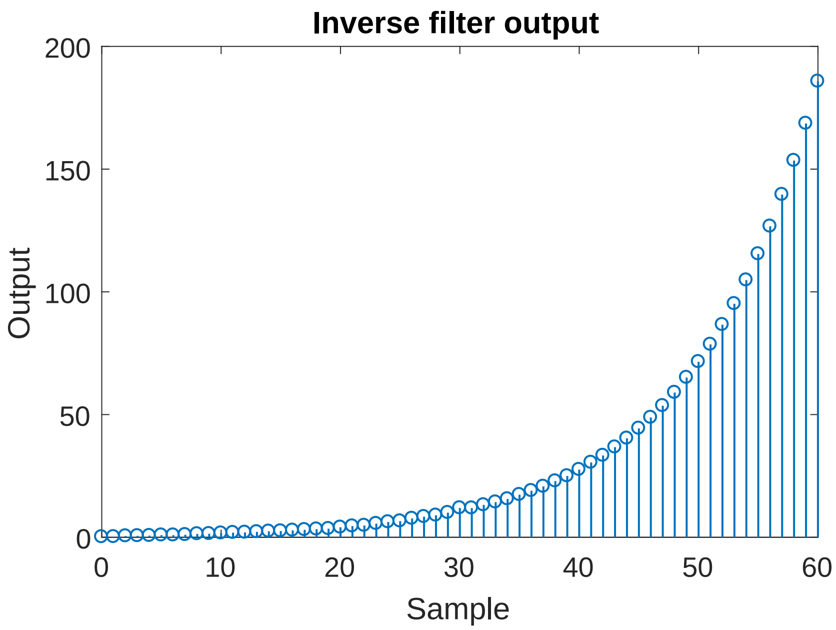
then $G(z)Y(z) = \frac{z^{-D}X(z)+N(z)}{X(z)} = z^{-D} + \frac{N(z)}{X(z)}$. The first term, z^{-D}, corresponds to a delta function in the time domain, indicating the presence of the signal at a given delay, D. The second is a filtered noise term.

Unfortunately, this process, whilst theoretically possible, is impractical due to noise, and quantisation errors. Should $X(z)$ contain any zeros outside the unit circle, then $G(z)$ will have poles outside the unit circle, and be unstable. Even if $X(z)$ was minimum phase (i.e. all zeros inside the unit circle), any zeros close to the unit circle will result in poles in $G(z)$ at the same location. Noise at frequencies close to this will be significantly amplified. Thus, inverse filtering is not appropriate in the general case.

Inverse filter example

Let $x(n) = \{1, -0.7, -0.3, -0.534, 0.6305, -0.23375\}$. $X(z)$ has five zeros, all inside the unit circle save for one zero located at $z = 1.1$. The corresponding inverse filter, $G(z) = \frac{1}{X(z)}$ has five poles, one of which is located at $z = 1.1$. Conceptually:





Right from the beginning of the output, the inverse filter exhibits problems, with a noise term that rapidly builds up, with an exponential rise. This is due to the presence of the pole of $G(z)$ that lies outside the unit circle. The delay at which the desired signal is present is at sample 30 corresponding to the signal delay. It is evident that the peak is lost in the exponential growth.

Obviously, an inverse filter cannot be applied to problems where the z-transform of the signal has zeros outside of the unit circle. Even where all the zeros lie inside the unit circle, imperfect cancellation can result in noise dominating the system output. Thus, in general, an inverse filter is generally not a good method to choose.

Cauchy-Schwartz Inequality

In order to develop a better technique, we start with the dot product of two real vectors:

$$\mathbf{x} \cdot \mathbf{y} = |\mathbf{x}||\mathbf{y}|\cos(\theta)$$

where θ is the angle between the vectors. From this it is clear that

$$\begin{aligned} |\mathbf{x} \cdot \mathbf{y}| &\leq |\mathbf{x}||\mathbf{y}| \\ \Rightarrow |\mathbf{x} \cdot \mathbf{y}|^2 &\leq |\mathbf{x}|^2|\mathbf{y}|^2 \end{aligned}$$

with equality holding only when $\mathbf{x} = k\mathbf{y}$. For a time series, $\mathbf{x} = [x(0)x(1)x(2)\ldots]$, and similarly for $\mathbf{y}$. Thus

$$\mathbf{x} \cdot \mathbf{y} = \sum_i x(i)y(i)$$

and

$$\left(\sum_{i=0}^{N-1} x(i)y(i)\right)^2 \leq \left(\sum_{i=0}^{N-1} (x(i))^2\right)\left(\sum_{i=0}^{N-1} (y(i))^2\right)$$

Where x and y are continuous, the summation can be replaced by integration. Equality holds only when $\mathbf{x} = k\mathbf{y}$ and k is a real scalar. This is the Cauchy-Schwartz Inequality that we will apply to the problem of matched filtering.

As above, we define a system such that $Y(z) = X(z) + N(z)$, and we wish to develop a $G(z)$ that, in some sense, detects the presence of $X(z)$ at some delay. (We know, from above, that an inverse filter is not suitable, so here we are looking for an alternative approach). If, instead of trying to reverse the process of $X(z)$ directly, we view the problem as one of maximising the signal to noise ratio for a given sample time, n_m, then we develop a matched filter. (Note we have yet to define n_m, but will do so later in the derivation).

We start by defining the component of the filter output due to the signal $x(n)$ by the Inverse Fourier transform at the detection delay, n_m:

$$\text{Signal} = \frac{1}{2\pi}\int_{-\pi}^{\pi} G(\omega)X(\omega)e^{j\omega n_m}d\omega$$

Using our knowledge of random signals, for the noise component we can write:

$$\text{Noise Power} = \frac{1}{2\pi}\int_{-\pi}^{\pi} |G(\omega)|^2\Gamma_{nn}(\omega)d\omega$$

If we assume that the noise is white, then $\Gamma_{nn}(\omega) = \sigma^2$.

Now, we can write the Signal to Noise Ratio (SNR), that is the signal power over the noise power, as:

$$\begin{aligned} \text{SNR} &= \frac{\left|\frac{1}{2\pi}\int_{-\pi}^{\pi} G(\omega)X(\omega)e^{j\omega n_m}d\omega\right|^2}{\frac{\sigma^2}{2\pi}\int_{-\pi}^{\pi}|G(\omega)|^2d\omega} \\ &= \frac{\left|\int_{-\pi}^{\pi} G(\omega)X(\omega)e^{j\omega n_m}d\omega\right|^2}{2\pi\sigma^2\int_{-\pi}^{\pi}|G(\omega)|^2d\omega} \end{aligned}$$

In order to maximise this at sample n_m, we need to select a suitable $G(\omega)$. Applying the Cauchy-Schwartz inequality to the numerator:

$$\left|\int_{-\pi}^{\pi} G(\omega)X(\omega)e^{j\omega n_m}d\omega\right|^2 \leq \int_{-\pi}^{\pi}|G(\omega)|^2\,d\omega\int_{-\pi}^{\pi}\left|X(\omega)e^{j\omega n_m}\right|^2d\omega$$

and thus,

$$\text{SNR} \leq \frac{\int_{-\pi}^{\pi}|X(\omega)|^2\,d\omega}{2\pi\sigma^2}$$

with equality holding when $G(\omega) = cX^*(\omega)e^{-j\omega n_m}$ for any non-zero scalar, c.

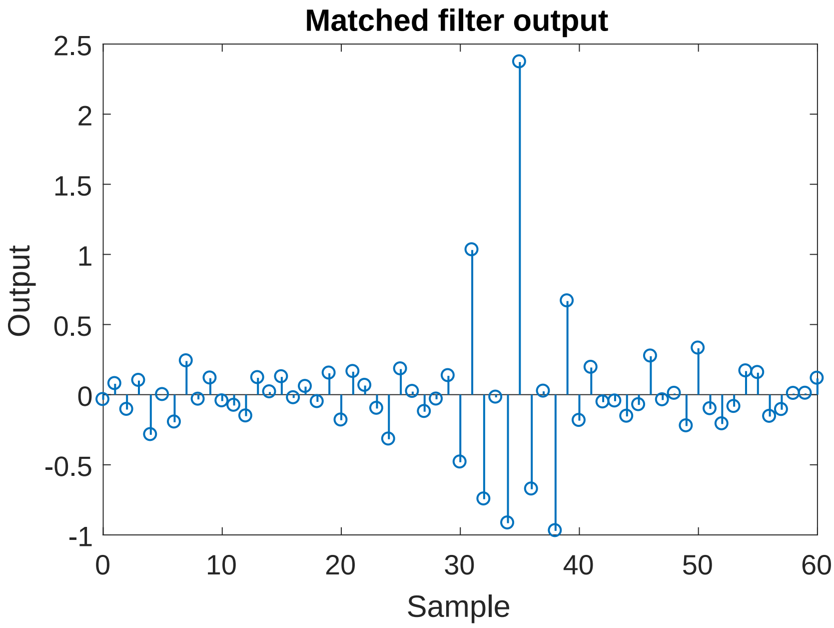
As the choice of c is arbitrary, and has no effect on the final SNR, we set $c = 1$. Thus,

$$\begin{aligned} G(\omega) &= X^*(\omega)e^{-j\omega n_m} \\ \Rightarrow g(n) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} X^*(\omega)e^{-j\omega n_m}e^{j\omega n}d\omega \\ &= \left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} X(\omega)e^{j\omega(n_m-n)}d\omega\right\}^* \\ &= \{x(n_m - n)\}^* \end{aligned}$$

If $x(n)$ is real, then $g(n) = x(n_m - n)$

Thus, the ideal filter impulse response is the complex conjugate of the time reversed sequence to be detected. We have yet to select the value of n_m, however it is clear that it should be chosen to make $g(n)$ causal. It also affects the time at which the maximum is detected, so the ideal n_m is equal to the length of the sequence $x(n)$.

For the same signal, the matched filter output is



The output of the filter, $w(n)$, is:

$$\begin{aligned}
w(n) &= \sum_{m=0}^{\infty} g(m) y(n-m) \\
&= \sum_{m=0}^{\infty} x(n_m - m) y(n-m) \\
&= \sum_{u=-\infty}^{n} x(n_m - n + u) y(u) \\
&= r_{yx}(n - n_m) = r_{xx}(n - D - n_m) + r_{nx}(n - n_m)
\end{aligned}$$

This result means that a matched filter is equivalent to performing a correlation with the signal to be detected. This implies that correlation, as introduced in the Discrete-Time Signals and Systems module, is the optimum technique for maximising detection of a known signal, from an SNR perspective.