The Cross-Correlation and Wiener-Khinchin theorems

The Cross-Correlation theorem relates the Fourier transform of the cross-correlation function of two signals with the Fourier transforms of the individual signals. Let $x(u)$ and $y(u)$ be two integrable signals, 0 everywhere outside $u \in [0, T)$, with values in $[-1, 1]$. The classic cross-correlation of these two signals is:

$$C_{xy}(\tau) = \int_{-\infty}^{\infty} x(u-\tau)y(u)du$$

where τ is the lag.[1]

The cross-correlation theorem states that:

$$C_{xy}(\tau) \iff \overline{X(f)}Y(f) \tag{1}$$

where $X(f)$ is the Fourier transform of $x(t)$ (which we denote $X(f) \iff x(t)$) and $\bar{\cdot}$ represents the complex conjugate.

The Wiener-Khinchin theorem is the special case where $x(t) = y(t)$, and substituting this into equation 1, we obtain:

$$C_{xx}(\tau) \iff |X(f)|^2 \tag{2}$$

Application to interaural time differences

Licklider (1959) demonstrated that a coincidence detector neuron could be treated as computing the cross-correlation between its two inputs. In the case of the interaural time difference (ITD), the two inputs are the signals presented to the two ears after filtering by the cochlea and the auditory system preceding the locus of ITD computation. Here, we assume that the two inputs are identical up to a time shift corresponding to the characteristic delay (CD) of the neuron, which we will treat as 0 without loss of generality. Thus, we expect the Wiener-Khinchin theorem (equation 2) to hold.

We would like to relate this to the *two-sided power spectral density* (PSD) of the input $x(t)$, which is given by:

$$\text{PSD}_x(f) = \frac{|X(f)|^2 + |X(-f)|^2}{2}.$$

Since $x(t)$ is real-valued, $X(f) = X(-f)$, and we have

$$\text{PSD}_x(f) = |X(f)|^2.$$

Therefore, by Equation 2, and letting ITD $= C_{xx}$, we have:

$$\begin{aligned} \text{PSD}_{\text{ITD}}(f) &= |(|X(f)|^2)|^2 \\ &= |\text{PSD}_x(f)|^2 \\ &= (\text{PSD}_x(f))^2. \end{aligned}$$

(since the PSD is always real and postive). Thus, the PSD of the autocorrelation of a signal is given by the square of the PSD of the signal itself. If we express power in decibels, which is customary, the *logarithmic two-sided power spectral density* (lPSD) is given by:

$$\begin{aligned} \text{lPSD}_{\text{ITD}}(f) &= 10\log_{10}(\text{PSD}_{\text{ITD}}(f)) \\ &= 10\log_{10}\left(\text{PSD}_x(f)^2\right) \\ &= 20\log_{10}(\text{PSD}_x(f)) \\ &= 2\text{lPSD}_x(f) \end{aligned}$$

So if the ITD can truly be treated as a cross-correlation, we expect the lPSD of the rate-ITD function to be double the lPSD of the effective input to the neuron.

[1] The conditions placed on x and y are not necessary for the definition of cross-correlation. However, they are necessary for later parts of the derivation, and match the nature of the signals used in the context of this research.

The Cross-Correlation theorem and running cross-correlation

The above relationship between the lPSD of the rate-ITD function and the lPSD of the effective input to the neuron was derived assuming that the coincidence detector response is the cross-correlation of the input stimuli. However, what Licklider (1959) demonstrated was that the instantaneous firing rate of a coincidence detector would be given by the *running cross-correlation* of the two stimuli. Choose a function $G(u)$ such that as $t \to \infty$, $G(u) \to 0$, and $G(u) = 0$, $\forall u < 0$. Then, the running cross-correlation is given by:

$$RC_{xy}(t, \tau) = \int_{-\infty}^{\infty} x(u - \tau) y(u) G(t - u) du.$$

The running cross-correlation is a function of both time and lag, and is therefore distinct from the cross-correlation. In particular, no result equating the running cross-correlation to any function of the Fourier transform exists. Therefore, to apply the Cross-Correlation theorem to neural data, it is necessary to first demonstrate the conditions under which running cross-correlation data can be treated as a classic cross-correlation.

In analyzing our data, we consider the mean number of spikes produced per stimulus. This is equivalent to averaging across time, and the *time-averaged running cross-correlation* is then:

$$\begin{aligned} TARC_{xy}(\tau) &= \frac{1}{T} \int_0^T RC_{xy}(t, \tau) dt \\ &= \frac{1}{T} \int_0^T \int_{-\infty}^{\infty} x(u - \tau) y(u) G(t - u) du\, dt \\ &= \frac{1}{T} \int_{-\infty}^{\infty} x(u - \tau) y(u) \left[\int_u^T G(t - u) dt \right] du. \end{aligned}$$

Let $G(u) = H'(u)$ for $u > 0$. Then,

$$\begin{aligned} TARC_{xy}(\tau) &= \frac{1}{T} \int_{-\infty}^{\infty} x(u - \tau) y(u) \left[H(T - u) - H(0) \right] du \\ &= \frac{1}{T} \left[-H(0) \int_{-\infty}^{\infty} x(u - \tau) y(u) du + \int_{-\infty}^{\infty} x(u - \tau) y(u) H(T - u) du \right] \\ &= \frac{1}{T} \left[-H(0) C_{xy}(\tau) + \int_{-\infty}^{\infty} x(u - \tau) y(u) H(T - u) du \right]. \end{aligned}$$

Note that since $y(u) = 0$ when $u > T$, we need not be concerned by the fact that $H(T - u)$ is not defined when $u > T$.

From the above, we conclude that the time-averaged running cross-correlation is simply a scaled version of the normal running cross-correlation, plus an error term. In the context of neural processing, $G(u)$ is generally assumed to be related to the membrane time constant, and of exponential form Yin et al. (1987). Thus, we examine this error term in the case where $G(u) = e^{-u/k}$. If that is the case, then $H(u) = -ke^{-u/k}$, and we have:

$$TARC_{xy}(\tau) = \frac{k}{T} \left[C_{xy}(\tau) - \int_0^T x(u - \tau) y(u) e^{\frac{u-T}{k}} du \right].$$

Since we assume that x and y have range $[-1, 1]$, we see:

$$\begin{aligned} \int_0^T x(u - \tau) y(u) e^{\frac{u-T}{k}} du &\leq \int_0^T e^{\frac{u-T}{k}} du \\ &\leq k e^{\frac{u-T}{k}} \Big|_0^T \\ &\leq k \left[1 - e^{\frac{-T}{k}} \right] \\ &< k. \end{aligned}$$

Similarly for -1, and so we have that the error term is in the range $(-k, k)$. However, it is not true that for all values of τ the absolute value of the error term will be less than $|C_{xy}|$. To see this, consider the case where $x(u-\tau)y(u)$ is the function $S(u)$:

$$S(u) \quad = \quad \begin{cases} -1 & u < T/2 \\ 1 & u \geq T/2 \end{cases}$$

In this case, C_{xy} will be zero, and the error term will be $2k(e^{-\tau/k} - e^{(T/2-\tau)/k})$, and hence nonzero.

In the case of the stimuli used in this study, a 50 ms long segment has C_{xy} with mean absolute values in the range of 3.31 ± 0.50 for the raw stimuli. If the stimuli are first convolved with the reverse-correlation estimated impulse response, C_{xy} is in the range of 14.6 ± 3.9. Since k is expected to be related to the membrane time constant, and hence have a magnitude no greater than on the order of 10^{-2}, it is reasonable to expect that the error term is insignificant. Thus, the time-averaged response of a coincidence detector should grossly behave as classic cross-correlation, and the Cross-Correlation theorem applies.

References

Licklider JCR (1959). Three auditory theories. In S. Koch (Ed.), Psychology: A study of a science, pp. 41-144. New York, New York: McGraw-Hill.

Yin TCT, Chan JCK, and Carney LH (1987). Effects of interaural time delays of noise stimuli on low-frequency cells in the cat's inferior colliculus. III. Evidence for cross-correlation. *J Neurosci*, 58: 562-583.