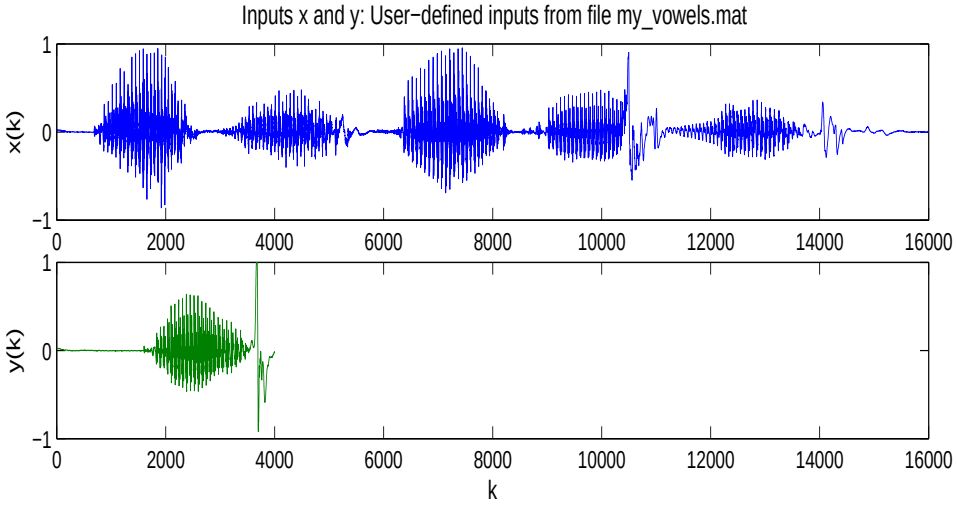
4.19 Use the GUI module *f_correlate* to record the sequence of vowels "A","E","I","O","U" in x. Play x to make sure you have a good recording of all five vowels. Then record the vowel "O" in y. Play y back to make sure you have a good recording of "O" that sounds similar to the "O" in x. Save this data in a MAT-file named *my_vowels*.

(a) Plot the inputs x and y showing the vowels.

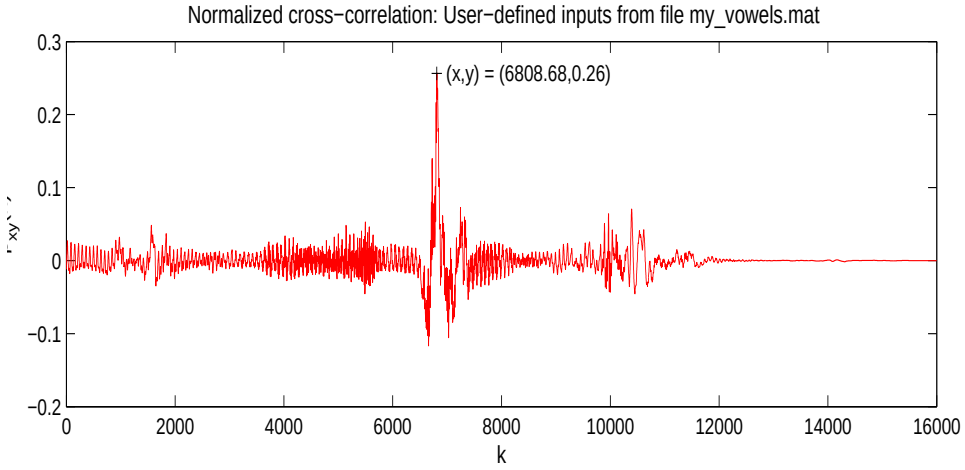
(b) Plot the normalized cross-correlation of x with y using the *Caliper* option to mark the peak which should show the location of y in x.

(c) Based on the plots in (a), estimate the lag d_1 that would be required to get the "O" in y to align with the "O" in x. Compare this with the peak location d_2 in (b). Find the percent error relative to the estimated lag d_1. There will be some error due to the overlap of y with adjacent vowels and coarticulation effects in creating x.

Solution



(a) The Vowels A, E, I, O, U



(b) Normalized Cross-correlation of x with y

(c) From part (a), the start of O in x is approximately $o_x = 9000$, and the start of O in y is approximately $o_y = 1700$. Thus the translation of y required to get a match with x is

$$\begin{aligned} d_1 &= o_x - o_y \\ &\approx 9000 - 1700 \\ &= 7300 \end{aligned}$$

The peak in part (b) is at $d_2 = 6809$. Thus the percent error in finding the location of O in x is

$$\begin{aligned} E &= \frac{100(d_2 - d_1)}{d_1} \\ &= \frac{100(7300 - 6809)}{7300} \\ &= 6.73\ \% \end{aligned}$$