Teoría de los filtros digitales para el procesamiento de señales.

Bronquivoide

Introducción:

Adjunto este PDF como una introducción a los notebooks sobre filtros FIR (*Finite-duration Impulse Response*), pues en comparación con otros notebooks, los fundamentos matemáticos son más extensos.

Normalmente adjunto la información necesaria y suficiente para comprender el contenido de las líneas de código, pero en esta ocasión, al ser más contenido, resultaría poco práctico atascar de información textual el jupyternotebook, pues al dedicar un notebook para cada filtro FIR, considero que lo más correcto es exponer los fundamentos del filtrado en un documento general, e ir especificando los casos particulares en su notebook correspondiente.

Y es más contenido teórico en esta parte de los filtros porque normalmente en el área de la física, de las matemáticas y la ingeniería, se cuentan con nociones sobre análisis real y complejo, sistemas continuos y discretos, transformadas integrales, funciones especiales, teoría de Sturm-Liouville de funciones ortogonales, estadística descriptiva, análisis de Fourier, etc (que son los temas que hemos abordado en los notebooks). Por ende, explicar todo desde cero no tiene mucho caso, sin embargo, con respecto a la teoría del filtrado considero que sí es necesario abordar los fundamentos más a profundidad.

Fundamentos matemáticos

En el notebook de convolución en señales habíamos hablado de que las señales descritas mediante un sistema LTI causal FIR seguían la siguiente forma de convolución:

$$y(n) = \sum_{k=0}^{N-1} x(k)h(n-k) \tag{1}$$

En donde M es el número de muestras de la señal analógica de entrada ya discretizada, $x(n)$; $y(n)$ es la señal de salida.

Partiendo de lo anterior, me gustaría abordar más a detalle, desde una perspectiva más formal lo correspondiente a un filtro asociado a esta clase de sistemas (filtro FIR).

Los filtros digitales tienen dos principales propósitos:

1. **Separación de señales:** Separar las señales por sus espectros de frecuencia (si las señales tienen diferentes dominios de frecuencias entonces consideramos que son espectralmente separables).

2. **Restauración de señales:** Restaurar una señal alterada, corrompida por ruido.

Una señal de tiempo discretizado es una secuencia de valores numéricos (pueden ser reales o complejos como vimos en el caso de la transformada de Fourier), la cual podemos expresar mediante una tupla infinita (pues puede ser un arreglo de longitud infinita), de la forma:

$$X = \begin{pmatrix} \vdots \\ x(-1) \\ x(0) \\ x(1) \\ \vdots \end{pmatrix} \tag{2}$$

En donde $x(n)$ es el valor de la señal a tiempo discretizado $n\tau$. Al arreglo X, dependiendo la literatura se le puede llamar de distintas maneras, como: vector señal, vector de observación, vector de muestreo digital, etc.

Sobre las bases matemáticas de X: Requerimos que sus componentes sean funciones escalares pertenecientes al espacio $\mathscr{L}_2$.

Recordando que el espacio continuo $\mathbf{L_2}$ puede estar definido sobre el campo de los reales o sobre el campo de los complejos como:

$$\mathbf{L_2}(\mathbb{K}) := \{f : \mathbb{K} \to \mathbb{K} | \int_{-\infty}^{\infty} |f(x)|^2 dx < \infty\}$$

En donde $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$, este tiene su espacio análogo discreto, $\mathscr{L}_2$, definido como:

$$\mathscr{L}_2(\mathbb{K}) := \{x : \mathbb{K} \to \mathbb{K} | \sum_{-\infty}^{\infty} |x(n)|^2 < \infty\}$$

$\mathscr{L}_2$ es un espacio de Hilbert con el siguiente producto interior definido:

$$\langle x_1(n), x_2(n) \rangle = \sum_{-\infty}^{\infty} x_1(n) x_2^*(n) \tag{3}$$

Definimos también los impulsos:

$$e_j(n) := \delta_{ij} = \begin{cases} 1 & \text{si } i = k \\ 0 & \text{si } i \neq k \end{cases} \tag{4}$$

En donde $j = 0, \pm 1, \pm 2, \pm 3, ...$(Nota: se utiliza esta notación en analogía a la comunmente utilizada en álgebra lineal). El conjunto $\{e_j(n)\}$ forma una base para el espacio de señales, bajo la relación de ortogonalidad:

$$\langle e_j(n), e_k(n) \rangle = \delta_{jk} \tag{5}$$

A tiempo $n = 0$, a $\delta = e(0)$ se le conoce como el impulso unitario.

Con base en lo anterior, podemos escribir a X representado en términos de la base:

$$X = \sum_{-\infty}^{\infty} x(n) e_n \tag{6}$$

En otros notebooks habíamos hablado del operador de convolución, y así como podemos ver la convolución como un operador, como un mapeo, podemos ver otras cantidades y procesos como operadores. Por ejemplo, la discretización de una señal puede verse como un proceso representado matemáticamente mediante un operador.

Dentro de este tema, el operador que nos interesa es el operador de filtro digital Φ. Previo a definir este operador vamos a abordar otros operadores que nos servirán como identidades más tarde.

Los procesos que mencionamos anteriormente en otros notebooks como el de retardo y adelanto temporal en una señal de tiempo discreto, de la forma $n \to n-k$, están representado mediante los siguientes operadores lineales:

- **Operador de retraso (delay operator o right shift):**

  $\mathbf{R} : \mathscr{L}_2 \to \mathscr{L}_2$

$x(n) \to x(n-1)$

Shiftea el tiempo discreto por una unidad temporal hacia la derecha.

- **Operador de adelanto (advance operator o left shift):**

  $\mathbf{L} : \mathscr{L}_2 \to \mathscr{L}_2$

  $x(n) \to x(n+1)$

  Shiftea el tiempo discreto por una unidad temporal hacia la izquierda.

Ahora, definimos el operador filtro digital como:

$$\Phi : \mathscr{L}_2 \to \mathscr{L}_2 \quad \text{, mediante} \quad \Phi X = Y \tag{7}$$

Es decir, hablamos de que Φ es una matriz que procesa la señal de entrada X y la convierte en una salida Y cuyas componentes pertenecen también al espacio $\mathscr{L}_2$.

Φ es invariante en el tiempo, es decir, $\Phi x(n-k) = y(n-k)$.

Las entradas de la matriz, $\Phi_{ij} = h(k)$, constituyen una matriz de Toeplitz.

En términos del operador $\mathbf{R}$, podemos escribir Φ con un shifteo de m unidades temporales como:

$$\Phi = \sum_{m=-\infty}^{\infty} h_m \mathbf{R}^m. \tag{8}$$

Como $\mathbf{R}^m e_j = e_{j+m}$ y considerando la ecuación 6, tenemos que:

$$\Phi X = \Phi(\sum_j x(j) e_j) = \sum_{m,j} x(j) h(m) e_{j+m} = \sum_{i,j} x(j) h(i-j) e_i = \sum_i y(i) e_i \tag{9}$$

Entonces

$$y(i) = \sum_j x(j) h(i-j) \tag{10}$$

Decimos que el filtro asociado es FIR si solo una cantidad finita de coeficientes $h(n)$ son distintos de cero.

El filtro es causal si $h(j) = 0$ para $j < 0$, y está completamente determinado por el vector de *impulse response*:

$$H = (h(0), h(1), ..., h(N)) \tag{11}$$

Notemos que la acción del filtro está dada por la convolución:

$$Y = \Phi X = H \diamond X \tag{12}$$

Filtrado

El filtrado consiste en cambiar la amplitud de las componentes de frecuencia de una señal, o eliminar (o atenuar lo más posible) para quedarse con el espectro deseado en un determinado experimento o trabajo.

A los sistemas LTI (*Linear Time Invariant*) que cambian la forma del espectro de frecuencias de una señal se les llama *Frequency-shaping Filters*.

A los sistemas diseñados para pasar únicamente algunas componentes de frecuencias sin distorción (idealmente) y para eliminar (o atenuar) las demás se les conoce como Filtros selectivos de frecuencia. Definimos el proceso de filtración mediante el uso de sistemas LTI con una apropiada respuesta de frecuencia escogida.

Bajo el contexto del dominio de frecuencia mencionado en cuanto a sistemas LTI, un sistema de este tipo va a estar caracterizado por una función que depende de la variable de frecuencia ω (respuesta de frecuencia). En caso de que sea confuso, la función de respuesta de frecuencia va a caracterizar al sistema LTI en el dominio de frecuencias.

Consideramos un sistema LTI inicialmente relajado, el cual es excitado con una señal de entrada arbitraria $x(n)$, entonces la respuesta va a estar dada por una señal de salida por la suma de convolución:

$$y(n) = \sum_{k=-\infty}^{\infty} h(k)x(n-k) \tag{13}$$

La ecuación anterior es la relación entrada-salida en el dominio temporal caracterizada por $h(n)$, por lo que para desarrollar una caracterización del sistema en el dominio de frecuencias, consideramos la señal de entrada:

$$x(n) = Ae^{i\omega n} \tag{14}$$

Entonces, de la ecuación 13 tenemos que:

$$y(n) = A \sum_{k=-\infty}^{\infty} h(k)e^{i\omega(n-k)} = Ae^{i\omega n} \sum_{k=-\infty}^{\infty} h(k)e^{-i\omega k} \tag{15}$$

Definimos la función:

$$H(\omega) := \sum_{k=-\infty}^{\infty} h(k)e^{-i\omega k} \tag{16}$$

$H(\omega)$ es la transformada de Fourier del pulso de muestreo unitario de respuesta, $h(k)$, del sistema. El conjunto $\{h(n)\}$ pertenece al espacio
$\mathscr{L}_1(\mathbb{R}) := \{h : \mathbb{R} \to \mathbb{R} | \sum_{-\infty}^{\infty} |h(n)| < \infty\}$.

De lo anterior, la ecuación 13 se escribe como:

$$y(n) = AH(\omega)e^{i\omega n} \tag{17}$$

Diseño de filtros

Hay cuatro tipos de filtros FIR digitales básicos:

- Filtro Pasa Bajas (*Lowpass Filter*).
- Filtro Pasa Altas (*Highpass Filter*).
- Filtro Pasa Bandas (*Bandpass Filter*).
- Filtro Frena Bandas (*Bandstop Filter*).

A estos se les llama filtros selectivos de frecuencia, el cual pasa exponenciales complejas que cumplen ciertas condiciones de frecuencias y rechaza cualquier otro tipo de frecuencias.

Diseñar un filtro digital consta de cuatro pasos:

1. **Definir el dominio de frecuencia.**

   Este paso depende del segundo, dependiendo el método se va a implementar el primer paso de una u otra forma.

   Si diseñamos un filtro por el método de ventada (*Window Method*), vamos a definir un dominio de frecuencia con frecuencias de corte.

Si diseñamos un filtro por el método de mínimos cuadrados, vamos a definir el dominio de frecuencia con frecuencias de corte y un ancho de transición.

2. **Generar el filtro:**

   Se genera el filtro en el dominio temporal por mínimos cuadrados o por el método de ventana.

3. **Evaluación del filtro:**

   Se evalua en el dominio de frecuencias usando una transformada de Fourier.

4. **Aplicación del filtro:**

   Se aplica el filtro evaluado para el procesamiento digital de una señal.

Sabemos que en el dominio de tiempo, una señal de salida es:

$$y(n) = x(n) \diamond h(n) \tag{18}$$

En donde $\diamond$ es el operador de convolución.

Y en el dominio de frecuencias tenemos una señal de salida:

$$Y(\omega) = X(\omega)H(\omega) \tag{19}$$

En donde $Y(\omega)$ es la función compleja (que puede tener módulo o fase) y $X(\omega)$ el espectro de la señal de entrada.

La ecuación (19) cumple las siguientes propiedades para su módulo y fase:

$$|Y(\omega)| = |X(\omega)| \cdot |H(\omega)| \tag{20}$$

$$\theta(Y) = \theta(H) + \theta(H) \tag{21}$$

En su forma más sencilla, un filtro pasa bajas posee una frecuencia de corte ω_c tal que un sistema LTI pasa exponenciales complejas $exp(i\omega t)$ para $-\omega_c \leq \omega \leq \omega_c$ y rechaza señales a cualquier otra frecuencia.

Un filtro pasa bajas ideal (sin cambio de fase) está caracterizado bajo:

$$H_{LP}(\omega) = \begin{cases} 1 & \text{si } |\omega| \leq \omega_c \\ 0 & \text{si } |\omega| > \omega_c \end{cases} \tag{22}$$

O equivalentemente, en el dominio temporal:

$$h_{LP}(n) = \frac{sen(\omega n)}{\pi n} \tag{23}$$

Si tenemos una fase lineal:

$$H_{LP}(\omega) = \begin{cases} exp(-i\omega n_\alpha) & \text{si } |\omega| \leq \omega_c \\ 0 & \text{si } \omega_c \leq |\omega| \leq \pi \end{cases} \tag{24}$$

O:

$$h_{LP}(n) = \frac{sen(\omega_c \cdot (n - n_\alpha))}{\pi(n - n_\alpha)} \tag{25}$$

Por otra parte, un filtro pasa altas ideal está caracterizado bajo:

$$H_{HP}(\omega) = \begin{cases} 0 & \text{si } |\omega| \leq \omega_c \\ 1 & \text{si } |\omega| > \omega_c \end{cases} \tag{26}$$

Fuentes consultadas y recomendadas:

1. Oppenheim, A. V., Willsky, A. S., & Nawab, S. H. (1996). Signals & systems. Prentice Hall.

2. Proakis, J. G., & Manolakis, D. K. (2007). Digital signal processing: Principles, algorithms, and applications (4th ed.). Pearson.

3. Damelin, S. B., & Miller, W. (2006). The mathematics of signal processing. Academic Press.

4. Smith, S. W. (2003). Digital signal processing: A practical guide for engineers and scientists. Newnes.