

Capítulo 2

Antecedentes

En este capítulo se explican las principales herramientas matemáticas en que se basa este trabajo. La primera sección se ocupa del controlador PID discreto, un controlador comúnmente empleado. La sección 2.2 trata sobre las funciones wavelet, funciones matemáticas que se emplean para el análisis de datos. La sección 2.3 explica las herramientas conocida como redes neuronales artificiales, construcciones matemáticas empleadas para aproximación y reconocimiento de funciones que pueden actualizarse mientras ocurren en el tiempo. En la última parte se explica la construcción de la red wavenet, combinando las dos herramientas anteriores.

2.1 Teoría de control PID

Un controlador PID es una unidad diseñada para crear un lazo de control estable y alcanzar el desempeño deseado mediante tres tipos de acciones básicas de corrección de error: proporcional, integral y derivativa, que dan origen al mnemónico PID. Cada una de estas acciones resuelve un problema específico relativo al error.

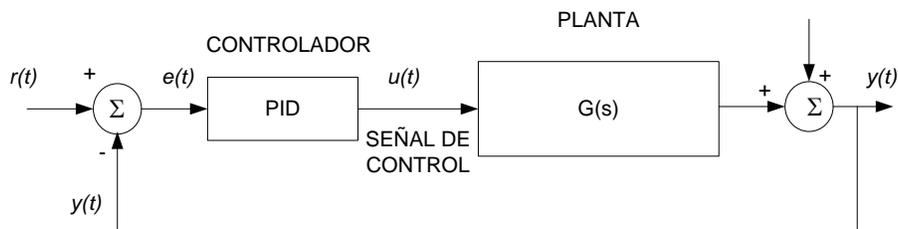


Figura 2.1. Esquema del controlador PID.

La acción proporcional es la principal respuesta del controlador ante las perturbaciones del lazo de control, ofrece una acción correctiva proporcional al tamaño del error. Si bien existe en todo el lazo de control, su presencia no garantiza que el cambio de valor de la

variable controlada coincide exactamente con el valor deseado para cualquier cambio de carga.

La función de la acción integral es eliminar el offset generado por la acción proporcional y asegurar que la salida del proceso corresponde a la señal de referencia. La acción proporcional genera un error en estado estacionario, pues una acción correctiva proporcional al tamaño del error no necesariamente conduce al valor de referencia.

La acción integral se agrega para asegurar que cuando el error acumulado es positivo se ofrece una señal de control que reduce el tamaño de la salida y cuando el error acumulado es negativo se ofrece una señal de control que aumenta la salida, sin importar qué tan pequeño es el error [5].

La acción derivativa mejora la velocidad de respuesta del controlador. Las anteriores acciones responden al error presente; la acción derivativa extrapola el error futuro como la tangente de la curva del error y hace la señal correctiva proporcional al error predicho.

El comportamiento del controlador PID es consecuencia de la aplicación paralela de estas tres acciones:

$$u(t) = K_c \left[e(t) + \frac{1}{T_i} \int_0^t e(t) dt + T_d \frac{de(t)}{dt} \right] \quad (2.1)$$

donde K_c es la ganancia proporcional, T_i el tiempo integral y T_d el tiempo derivativo.

A partir de la ecuación (2.1) se puede obtener la función de transferencia siguiente:

$$u(s) = K_c \left[1 + \frac{1}{T_i s} + T_d s \right] E(s) \quad (2.2)$$

Es posible observar que la ganancia proporcional actúa sobre el total de la respuesta. Las constantes integral y derivativa representan el peso relativo de cada una de estas acciones respecto a la acción proporcional, originando los siguientes parámetros:

$$K_i = \frac{K_c}{T_i}, \quad y \quad K_d = K_c T_d \quad (2.3)$$

2.2 Funciones wavelet

El término wavelet (ondeleta u ondita) se refiere a un grupo de funciones que presentan una pequeña oscilación y alcanzan rápidamente cero. Es posible emplear conjuntos de wavelets para aproximar una señal utilizando versiones trasladadas y dilatadas o contraídas en el tiempo para representar mejor la señal. De esta manera, son herramientas de descomposición de funciones que permiten el análisis en diferentes niveles de frecuencia y tiempo de una función. Una función wavelet es una función $\psi \in L^2$ que cumple con la siguiente condición de admisibilidad:

$$c_h = \int_{-\infty}^{\infty} \frac{|\Psi(\omega)|^2}{|\omega|} d\omega < \infty \tag{2.4}$$

donde $\Psi(\omega)$ es la transformada de Fourier de la función $\psi(t)$.

La función wavelet generalmente se conoce como wavelet madre, porque a partir de ésta se generan funciones “wavelet hijas” a partir de operaciones de dilatación y traslación,

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right), \quad a > 0; \quad a, b \in \mathbb{R} \tag{2.5}$$

donde $\psi_{a,b}(t)$ es la wavelet hija, a es el factor de dilatación-contracción, que varía la duración de la wavelet, y b es un desplazamiento en el tiempo. $\frac{1}{\sqrt{a}}$ es un factor de normalización que permite que

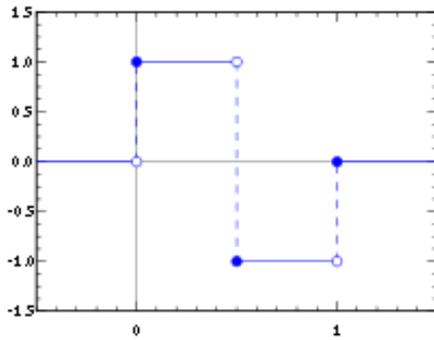
$$\|\psi_{a,b}\| = \|\psi\|, \quad \forall a, b \in \mathbb{R}.$$

En la Tabla 2.1 se presentan algunos ejemplos de wavelets comúnmente empleadas.

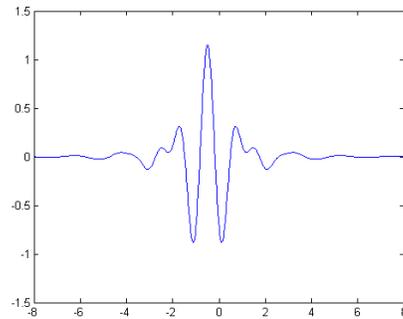
Wavelet	Función
Haar	$\psi(t) = \begin{cases} 1, & \text{si } 0 \leq t \leq \frac{1}{2}, \\ -1, & \text{si } \frac{1}{2} \leq t \leq 1, \\ 0, & \text{en cualquier otro caso} \end{cases}$
Meyer	$\psi(w) = \begin{cases} (2\pi)^{-\frac{1}{2}} e^{\frac{iw}{2}} \sin\left(\frac{\pi}{2} v \left(\frac{3}{2\pi} w - 1\right)\right), & \text{si } \frac{2\pi}{3} \leq w \leq \frac{4\pi}{3}, \\ (2\pi)^{-\frac{1}{2}} e^{\frac{iw}{2}} \cos\left(\frac{\pi}{2} v \left(\frac{3}{4\pi} w - 1\right)\right), & \text{si } \frac{4\pi}{3} \leq w \leq \frac{8\pi}{3}, \\ 0, & \text{si } w \notin \left[\frac{2\pi}{3}, \frac{8\pi}{3}\right] \end{cases}$
Morlet	$\psi(t) = e^{-\frac{t^2}{2}} \cos(5t)$

Mexican Hat	$\psi(t) = \frac{3}{\sqrt{3}} \pi^{-\frac{1}{4}} (1 - t^2) e^{-\frac{1}{2}t^2}$
RASP1	$\psi(t) = \frac{t}{t^2 + 1}$

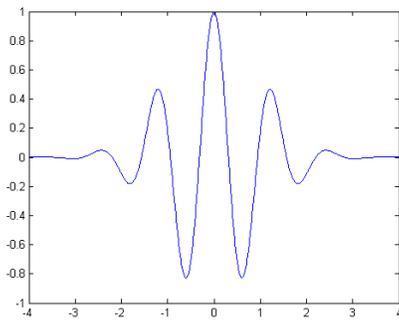
Tabla 2.1. Algunas wavelets madre.



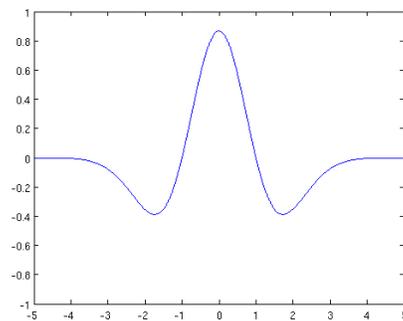
Wavelet Haar



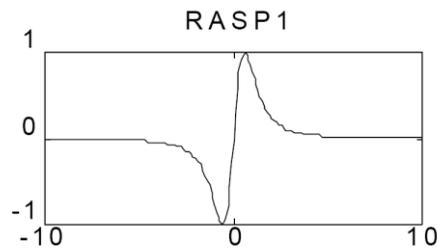
Wavelet Meyer



Wavelet Morlet



Wavelet Mexican Hat



Wavelet RASP1

Figura 2.2. Algunas wavelets comunes.

Las funciones wavelet descritas anteriormente forman la base de la transformada wavelet. La transformada wavelet es un operador que transforma una función integrándola con versiones modificadas de una función, en este caso, una wavelet madre.

La transformada wavelet de una función $f \in L^2$ respecto a una wavelet admisible está definida como:

$$W_f(a, b) = \int_{-\infty}^{\infty} f(t) \psi\left(\frac{t-b}{a}\right) dt$$

para $a \neq 0$, donde $\psi(\cdot)$ es la función wavelet. El efecto de traslación provoca que la transformada wavelet realice un efecto de ampliación en fenómenos de alta frecuencia y de breve duración. Conforme nos movemos de escalas mayores a menores, se pueden conseguir mejores representaciones de estos momentos en las señales.

2.3 Redes neuronales artificiales

Una red neuronal artificial es un modelo matemático adaptable que emula el funcionamiento de una neurona biológica por medio de la interconexión de nodos de proceso simple (neuronas artificiales).

La estructura de una red neuronal simple de L neuronas es la siguiente:

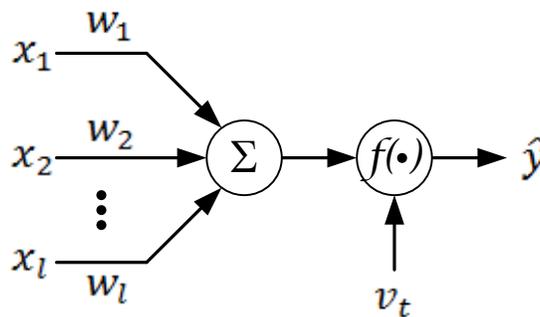


Figura 2.3. Estructura básica de una red neuronal.

donde x_l es la l -ésima entrada a la neurona, w_l es el peso asociado a dicha entrada, $f(\cdot)$ es la función de activación de la neurona, \hat{y} la señal de salida y v_t la señal de bias o umbral.

La función de salida de la neurona es función de la suma ponderada de las entradas:

$$\hat{y} = f\left(\sum_{l=1}^L w_l x_l - v_t\right) \quad (2.6)$$

Esto es importante porque podemos establecer una función de aprendizaje a partir del gradiente de \hat{y} para permitir que los pesos de cada conexión se actualicen.

Se define una función de medida del error E como:

$$E = \frac{1}{2}(y - \hat{y})^2 \quad (2.7)$$

donde y es la función que se desea aproximar o función objetivo y \hat{y} es la salida de la red neuronal.

De acuerdo con el método del gradiente inverso, obtenemos Δw como la derivada parcial del error respecto a w .

$$\Delta w(k) = -\frac{\partial E}{\partial w} \quad (2.8)$$

El cambio de w para la siguiente iteración se obtiene de la siguiente manera:

$$w(k+1) = w(k) + \mu_w \Delta w(k) \quad (2.9)$$

Donde μ_w es la velocidad de aprendizaje de la neurona.

2.4 Teoría wavenet

La teoría wavenet combina los dos conceptos vistos previamente, teoría de las redes neuronales artificiales y funciones wavelet para generar un algoritmo de aproximación de funciones. La figura 2.4 a continuación muestra la arquitectura básica de una red wavenet adaptable:

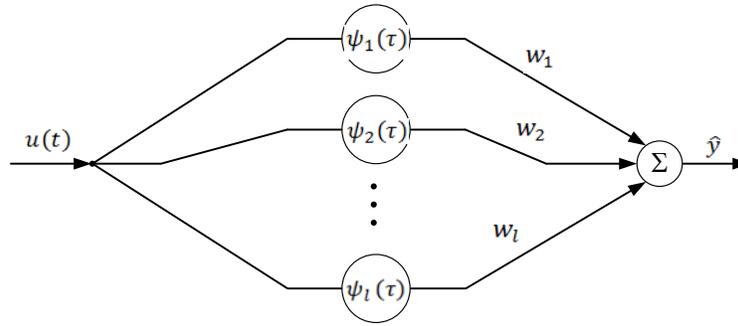


Figura 2.4. Diagrama de una red neuronal wavenet.

En la red wavenet $u(t)$ es la señal de entrada, $\hat{y}(t)$ la señal de salida de la red, $\psi_l(\tau)$ es la función wavelet de activación de la l -ésima neurona y w_l es el peso de conexión entre las neuronas.

Una aproximación razonable se consigue si las regiones de tiempo de la señal están cubiertas adecuadamente por las L ventanas de la red neuronal. La señal aproximada por la red es

$$\hat{y}(t) = u(t) \sum_{l=1}^L w_l \psi_l(\tau), \quad \hat{y}, u, w \in \mathbb{R} \tag{2.10}$$

Se definen los siguientes vectores:

$$\mathbf{A}(k) \triangleq [a_1(k) \quad a_2(k) \quad \dots \quad a_l(k) \quad \dots \quad a_{L-1}(k) \quad a_L(k)]^T \tag{2.11}$$

$$\mathbf{B}(k) \triangleq [b_1(k) \quad b_2(k) \quad \dots \quad b_l(k) \quad \dots \quad b_{L-1}(k) \quad b_L(k)]^T \tag{2.12}$$

$$\mathbf{W}(k) \triangleq [w_1(k) \quad w_2(k) \quad \dots \quad w_l(k) \quad \dots \quad w_{L-1}(k) \quad w_L(k)]^T \tag{2.13}$$

$$\mathbf{\Psi}(\tau) \triangleq [\psi_1(\tau) \quad \psi_2(\tau) \quad \dots \quad \psi_l(\tau) \quad \dots \quad \psi_{L-1}(\tau) \quad \psi_L(\tau)]^T \tag{2.14}$$

Y con ello es posible representar la ecuación (2.10) como

$$\hat{y} = u(t) \mathbf{\Psi}^T(\tau) \mathbf{W}(k) \tag{2.15}$$

Adicionalmente, se puede lograr una aproximación doble colocando un filtro de respuesta infinita al impulso en cascada con la red neuronal. El filtro IIR se emplea como un segundo nivel de identificación y aproximación de la señal.

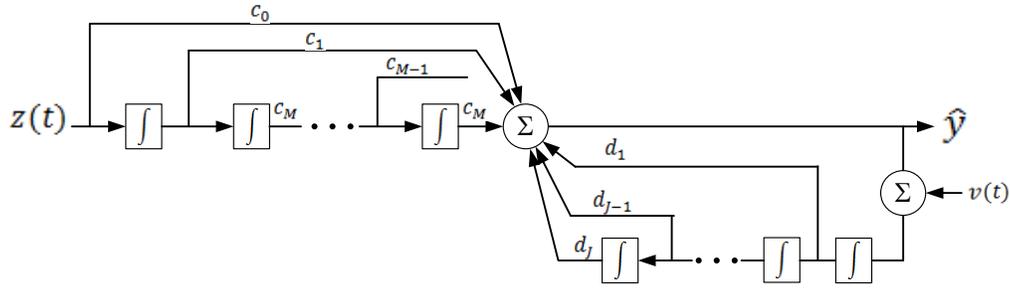


Figura 2.5. Diagrama del filtro IIR. $z(t)$ es la señal de entrada, $\hat{y}(t)$ es la señal de salida, c y d son los coeficientes del filtro, $v(t)$ es la señal de bias.

La estructura definitiva de la red se muestra a continuación:

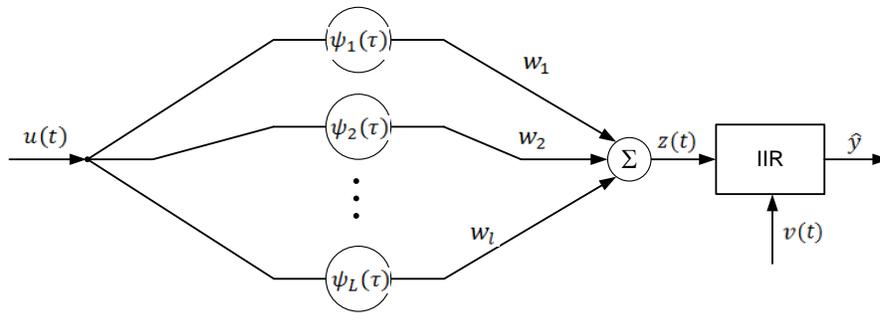


Figura 2.6. Red neuronal en cascada con un filtro IIR.

Se definen los vectores \mathbf{C} y \mathbf{D} de coeficientes del filtro, el vector \mathbf{Z} que contiene las salidas actual y anteriores de la red neuronal y el vector $\hat{\mathbf{Y}}$ que contiene las salidas actual y anteriores del filtro IIR.

$$\mathbf{C}(k) \triangleq [c_1(k) \quad c_2(k) \quad \dots \quad c_m(k) \quad \dots \quad c_{M-1}(k) \quad c_M(k)]^T \quad (2.16)$$

$$\mathbf{D}(k) \triangleq [d_1(k) \quad d_2(k) \quad \dots \quad d_j(k) \quad \dots \quad d_{j-1}(k) \quad d_j(k)]^T \quad (2.17)$$

$$\mathbf{Z}(k) \triangleq [z(k) \quad z(k-1) \quad \dots \quad z(k-m) \quad \dots \quad z(k-M+1) \quad z(k-M)]^T \quad (2.18)$$

$$\hat{\mathbf{Y}}(k) \triangleq [\hat{y}(k) \quad \hat{y}(k-1) \quad \dots \quad \hat{y}(k-j) \quad \dots \quad \hat{y}(k-J+1) \quad \hat{y}(k-J)]^T \quad (2.19)$$

donde

$$z(k) = \Psi^T(\tau)\mathbf{W}(k) \quad (2.20)$$

La función de salida de la red, la función de aproximación \hat{y} se puede expresar de la siguiente forma:

$$\hat{y} = u(k)\mathbf{C}^T(k)\mathbf{Z}(k) + v(k)\mathbf{D}^T(k)\hat{\mathbf{Y}}(k) \quad (2.21)$$

Es posible actualizar los parámetros de la red wavenet para reducir el error entre la estimación y la señal deseada, mejorando la señal de aproximación \hat{y} . Se emplea un

algoritmo de mínimos cuadrados medios por medio de la minimización de la función de costo o función de error E (que es una función de energía) a lo largo de todo el tiempo.

Los coeficientes son actualizados como se muestra en la ecuación (2.9) de acuerdo con la siguiente ley:

$$\begin{aligned}\mathbf{A}(k + 1) &= \mathbf{A}(k) + \mu_A \Delta \mathbf{A} \\ \mathbf{B}(k + 1) &= \mathbf{B}(k) + \mu_B \Delta \mathbf{B} \\ \mathbf{W}(k + 1) &= \mathbf{W}(k) + \mu_W \Delta \mathbf{W} \\ \mathbf{C}(k + 1) &= \mathbf{C}(k) + \mu_C \Delta \mathbf{C} \\ \mathbf{D}(k + 1) &= \mathbf{D}(k) + \mu_D \Delta \mathbf{D}\end{aligned}\tag{2.22}$$

μ es la velocidad de aprendizaje de cada parámetro.