

Resumen “Procesamiento de señales”

Grupo de la materia: <https://groups.google.com/forum/#!forum/psfrbautn>

Sitio de la materia: <https://sites.google.com/site/psfrbautn/home/clases>

Cartas sobre Estadística

Estadística Descriptiva, Representación Gráfica y Distribución Normal.

Estadística descriptiva

Una definición de “estadística descriptiva” es “describir los datos en forma concisa”, la forma más común de describir un conjunto de datos relacionados entre sí es reportar un valor medio y una dispersión alrededor de dicho valor medio.

La medida básica para describir el valor central de un conjunto de datos es el valor medio o media del mismo. La Segunda medida de valor “central” de un conjunto de datos es la mediana, definida como el valor en el medio cuando los datos son ordenados de menor a mayor. La media geométrica es otra forma de describir el valor central de un conjunto de datos, es en realidad una forma transformada de calcular la media, en ciertos casos como por ejemplo un conjunto de datos que proviene de mediciones múltiplo de 2, el valor medio no describe adecuadamente el valor central. La moda es otra forma de describir el valor central de un conjunto de datos, se calcula como el valor más frecuente. En el caso que ningún dato del conjunto tenga una frecuencia mayor a 1, el resultado de la moda es nulo.

Para poder describir mejor un conjunto de datos necesitamos una medida de dispersión además de una del valor central, la más simple es el rango, el cual muestra los valores mínimo y máximo del conjunto de datos. La varianza y el desvío estándar son las medidas de dispersión más populares. El desvío estándar se define como la raíz cuadrada de la varianza. Otras medidas de dispersión más sofisticadas pero no menos útiles son:

- cuantiles
- percentilos: se utiliza cuando la probabilidad tiene un valor entre 0% y 100%, siendo los valores mínimo y máximo del conjunto de datos respectivamente.

El percentilo 10to se interpreta entonces como el valor para el cual el 10% de los datos del conjunto son menores al mismo. Otra medida de dispersión son los cinco números de Tukey que calcula cinco valores que describen concisamente un conjunto de datos, son los valores mínimo, los percentilos 25to, 50to y 75to, y el valor máximo.

Otra medida de dispersión es el coeficiente de variación que es el cociente del desvío estándar y la media expresado en %, esta medida es fácil de interpretar, pero en la mayoría de los casos no es adecuada para comparar dos más conjuntos de datos.

Representación gráfica

El gráfico de caja (box-plot en Inglés) es la forma gráfica de los cinco números, como podemos ver en la figura 1 la caja muestra los percentilos 25to y 75to, la línea en el medio de la caja es la mediana (percentilo 50to), los extremos muestran los valores mínimo y máximo. Otra opción, muy utilizada por estadísticos, es el gráfico de rama y hoja (stem-and-leaf plot en Inglés):

2 | 2234

2 | 5567899999

Los números que se muestran a la izquierda del carácter | son los dígitos más significativos, el punto decimal está ubicado un dígito a la derecha del carácter |, en otras palabras la primera línea 2 | 2234 se lee como el primer valor 22 (por el 2 | 2), luego hay otro 22, un 23 y un 24, correspondiente a los tres siguientes números, todos ellos corresponden a los sujetos 1 a 4.

La representación gráfica más popular de un conjunto de datos es el histograma, el cual representa la frecuencia de aparición de valores dentro del rango del conjunto de datos.

Distribución normal

La distribución normal se define con su función de densidad de probabilidad.

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right], \text{ para } -\infty < x < \infty$$

donde los parámetros μ y σ son la media y el desvío estándar respectivamente. Por definición para $-\sigma < x < \sigma$ el área bajo la curva es el 68% del área total, para

$-1.96\sigma < x < 1.96\sigma$ es el 95% del área, y para $-2.576\sigma < x < 2.576\sigma$ el 99% del área total,

Unas muestras de una misma población normal pueden ser diferentes, y por lo tanto la media de las mismas también será diferente, y dichas medias pueden también tener una distribución. Un teorema fundamental de la estadística dice que las medias de muestras aleatorias provenientes de cualquier distribución tiene una distribución normal, dicho teorema se conoce con el nombre de “teorema del límite central”, una consecuencia de este teorema es que cuando trabajamos con muestras de cientos de observaciones podemos olvidarnos de la distribución de la población y asumir que es normal. Una regla práctica muy utilizada dice que muestras con 30 o más observaciones tienen una distribución aceptablemente normal.

Una de las aplicaciones más importantes del teorema del límite central es la posibilidad de calcular los denominados intervalos de confianza (IC), el más utilizado es el IC del 95%, por ejemplo si conocemos la media (μ) el desvío estándar (σ) de una muestra por definición el 95% de los datos se encuentran dentro del intervalo determinado por

$\mu - 1.96\sigma$ y $\mu + 1.96\sigma$.

Prueba de hipótesis y de tendencia central sobre una y dos muestras

Formulación y prueba de hipótesis

Los pasos del método científico se pueden resumir de la siguiente forma:

- 1) plantear el problema a resolver,
- 2) efectuar las observaciones,
- 3) formular una o más hipótesis,
- 4) probar dichas hipótesis
- 5) proclamar las conclusiones;

La estadística nos puede ayudar en los pasos 2) (diseño de las observaciones) y 4) (prueba de hipótesis).

Una definición de hipótesis es la siguiente: “una explicación tentativa que cuenta con un conjunto de hechos y puede ser probada con una investigación posterior”.

La formulación formal de una hipótesis en el método científico se realiza definiendo la hipótesis nula (H_0) y la hipótesis alternativa (H_1); generalmente la H_0 establece que no hay diferencias entre el grupo control y el grupo investigado. La hipótesis alternativa (H_1) por otra parte, suele indicarse como el complemento de la H_0 , por lo tanto la H_1 . Por lo tanto la prueba de nuestra hipótesis consistiría en arbitrar los procedimientos necesarios para intentar rechazar o aceptar la hipótesis nula dependiendo el caso de estudio.

A la hora de tomar una decisión respecto de la hipótesis nula, surgen situaciones que nos pueden llevar a cometer diferentes errores. Si tomamos como ejemplo que lo que se busca es rechazar la hipótesis nula, puede que lleguemos a la conclusión de que el enunciado de nuestra H_0 sea verdadero en tal caso no rechazamos nuestra H_0 o bien que sea falso, en cuyo caso rechazaremos la H_0 . En esta situación puede que hayamos rechazado la H_0 cuando en realidad era cierta, o que la evidencia no haya sido suficiente para rechazarla siendo falsa. Estas diferentes situaciones plantean la existencia de diferentes tipos de errores

- Error tipo I, también denominado error α , se produce cuando **se rechazó la H_0 y es verdadera**. Éste, representa la probabilidad de haber cometido este tipo de error. Se

establece a priori α como el nivel de significancia o error máximo aceptable para la conclusión.

- **Error tipo II**, o también denominado error β , se produce cuando **se acepta la H_0 siendo falsa**. El error de tipo II está asociado con la potencia del método estadístico utilizado para poder detectar diferencias. La potencia de un método estadístico en una determinada situación se calcula como $(1 - \beta)$, lo que se corresponde con la situación de haber rechazado correctamente la H_0 , pues era falsa. Al igual que el valor de significancia α , la potencia del método estadístico se establece por el tamaño de la muestra y la prueba estadística utilizada.

Una hipótesis no se acepta, simplemente la evidencia no alcanza para rechazarla, y se mantiene como cierta mientras no se rechace.

En cualquier caso rechazar la H_0 es lo mismo que aceptar la H_1 y viceversa. El resultado final de un método estadístico para la prueba de una hipótesis es el valor P , que indica la probabilidad de obtener un valor más extremo que el observado si la H_0 es verdadera. Cuando P es menor que α se procede a rechazar la H_0 .

Pruebas de normalidad de una muestra

El paso inicial es determinar si las variables en estudio pueden ser representadas por una distribución normal. La importancia de verificar la normalidad de las muestras en estudio es fundamental en estadística porque si las muestras son normales se pueden aplicar métodos estadísticos paramétricos convencionales.

Los métodos de la estadística descriptiva nos pueden ayudar a verificar la normalidad de las variables, un histograma y un gráfico de cajas nos muestra en dos formas distintas la distribución de los datos. El cálculo de los cinco números de Tukey nos muestran numéricamente que la no hay evidencia suficiente como para rechazar la distribución normal de la variable IMC.

Unas pruebas de normalidad más formales, no paramétricas, muy recomendables para verificar la normalidad de una variable son las pruebas de Shapiro-Wilk, y de Kolmogorov-Smirnov.

Contrariamente a lo que se desea en la mayoría de los casos, en las pruebas de normalidad se busca aceptar la H_0 , dado que en la mayoría de los métodos estadísticos convencionales es necesaria la distribución normal de la variable de interés, pues siendo así es posible conocer los parámetros que la describen por completo, su media (μ), su desvío (σ) y la relación entre ambos y en este sentido estos métodos son más potentes. Un valor $P \geq 0.05$ en los tests de normalidad indicaría que no hay prueba suficiente para rechazar la normalidad de la variable.

Pruebas sobre una muestra

Suponga para que consideramos al grupo 1, los valores para la población normal de una ciudad y los del grupo 2 una muestra. Los parámetros del grupo 1 son considerados los poblacionales y los del grupo 2 los de la muestra. En este caso existe un estadístico que permite comparar la media muestral con la población.

En ciertas ocasiones se cuenta con una muestra y con la media y el Desvío Estándar de una población, en dicho caso se utiliza la prueba sobre una muestra calculando el estadístico Z :

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

donde \bar{x} es la media de la muestra (grupo en estudio), μ_0 es la media la población, σ es el desvío estándar de la población, y n es el tamaño de la muestra.

Una vez obtenido el z se busca en la tabla de valores críticos de la distribución normal entre que niveles de significancia está, luego se lo compara con P y se rechaza o acepta la hipótesis según el caso de estudio. (recordar $P < 0.05$)

Pruebas sobre dos muestras (de una población específica)

Suponga ahora que los sujetos del grupo 1 y 2 corresponden ambos a muestras de una supuesta población subyacente. El test implicado intentará probar si ambas medias no difieren, lo que implica que ambas muestras provienen de la misma población y contrariamente si difieren.

La prueba t es la prueba paramétrica más utilizada; la misma está basada en el cálculo del estadístico t y de los grados de libertad, con estos dos resultados y utilizando o bien una tabla o bien un cálculo de la distribución t se puede calcular el valor de P.

La prueba t del estudiante se basa en tres supuestos:

1. uno es el de la distribución normal de los errores
2. es la independencia de los mismos,
3. es el de la homogeneidad de varianzas, considerado como el supuesto más importante.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad gl = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{s_1^4}{n_1^2(n_1-1)} + \frac{s_2^4}{n_2^2(n_2-1)}}$$

donde \bar{x}_1 y \bar{x}_2 son las medias de cada muestra (grupos); s_1^2 y s_2^2 son las varianzas de las muestras; n_1 y n_2 son los tamaños de la muestras.

Pruebas sobre dos muestras apareadas (misma población en diferentes momentos)

En ciertas ocasiones necesitamos trabajar sobre un mismo grupo de sujetos al cual se los observa en forma repetida, por ejemplo antes y después de un tratamiento, en este caso los sujetos son controles de ellos mismos. La prueba t es distinta para poder tener en cuenta que las observaciones son repetidas sobre el mismo grupo de sujetos. El primer paso es calcular el desvío estándar de las diferencia con la siguiente ecuación:

$$s = \sqrt{\frac{\sum_{i=1}^n (d_i - \bar{d})^2}{n-1}}$$

donde d_i es la diferencia entre dos mediciones consecutivas para cada sujeto; \bar{d} es la media de las diferencias; n es la cantidad de pares de observaciones.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s/\sqrt{n}}$$

y gl = n-1

Pruebas sobre dos muestras no normales

Siempre que se pueda es aconsejable transformar la muestra para que sea de distribución normal y así poder aplicar los métodos clásicos.

La transformación de la cual estamos hablando es numérica, puede ser simplemente calcular el logaritmo natural de cada observación, y luego verificar la normalidad de la muestra transformada. Por lo tanto el test me dirá si los logaritmos de las variables difieren o no, en este caso se debería considerar si esto tiene interpretación en el caso de estudio planteado.

Alternativa a la prueba de hipótesis

Una alternativa a las pruebas de hipótesis es el cálculo del intervalo de confianza del 95% de la media (IC 95% de la media), también denominado de la verdadera media, siempre que sea posible.

$$IC95\%media = \bar{x} \pm t_{gl,0.05} \frac{s}{\sqrt{n}}$$

donde \bar{x} , s y n son la media, el desvío estándar y el tamaño de la muestra respectivamente; $t_{gl,0.05}$ es el valor crítico de la distribución t para un determinado grado de libertad $gl=n-1$ y $\alpha=0.05$.

Por supuesto siempre es aconsejable verificar la normalidad de las variables ya que este método asume que las muestras son normales.

¿cómo se interpretan estos resultados ? La estimación del IC 95% de la media es un método alternativo a la prueba de hipótesis, la utilización del IC 95% de la media significa que la verdadera media puede tener por chance un valor dentro de dicho intervalo de confianza. Si los valores medios de cada grupo dentro de los respectivos intervalos de confianza no se solapan, podemos concluir diciendo que son distintos.

Regresión y Correlación

Las variables se pueden dividir en dos grupos:

- variable dependiente
- variables independientes

El análisis de la varianza puede ser considerado como un caso especial de la regresión.

Análisis de Regresión

El análisis de regresión se puede utilizar para describir la relación , su extensión, dirección e intensidad, entre una o varias variables independientes con escala intervalar y una variable dependiente también intervalar. Cabe destacar que no debe utilizarse el análisis de regresión como prueba de causalidad , en realidad no hay métodos estadísticos para probar causalidad.

La correcta aplicación del análisis de regresión asume que se cumplen las siguientes condiciones:

- Existencia:
 - Para cada valor fijo de X , existe un valor Y aleatorio con una distribución de probabilidad con valores finitos de media y varianza.
 - Esta condición debe cumplirse siempre;
- Independencia:
 - Los valores Y son estadísticamente independientes uno de otro
 - Esta condición puede no cumplirse en el caso de varias observaciones (valores de Y) sobre un mismo valor de X , en ese caso debe tenerse en cuenta dicha condición;
- Linealidad:
 - El valor medio de Y es una función lineal de X;
- Distribución normal:
 - Por cada valor fijo de X , Y tiene una distribución normal.

El analisis de regresion mas simple es el lineal y con una sola variable independiente, la variable independiente X se relaciona con la variable dependiente Y de acuerdo con la siguiente ecuación:

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

donde ε son los residuos, es decir la diferencia entre la estimación y los valores reales para cada par de puntos XY . Los coeficientes son: β_0 denominado intersección (cuando $X=0$) y β_1 pendiente.

Intervalos de confianza de la regresión

En algunas aplicaciones puede ser útil calcular el intervalo de confianza de la recta de regresión, de esta forma podemos verificar para una probabilidad dada (generalmente 95%), hasta donde se puede extender la recta de regresión. Por otro lado puede ser también conveniente calcular las bandas de predicción (BP).

Correlación

El complemento del análisis de regresión es el cálculo de la correlación, que se utiliza para cuantificar el grado de asociación de dos variables, note que en este caso no se considera a una variable dependiente y a la otra independiente, las dos tiene el mismo status. El coeficiente de correlación r , también denominado de Pearson se define como:

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

X e Y pueden intercambiar su lugar y el resultado es el mismo. El coeficiente de correlación puede tener valores dentro del siguiente rango: $-1 \leq r \leq 1$, el signo indica la dirección de la asociación, por lo tanto las asociaciones más fuertes son -1 y 1 , en el centro $r = 0$ indica falta de asociación. Otra medida derivada de r es el coeficiente de determinación, el cual simplemente se calcula como el cuadrado de r , por lo tanto el rango de valores es $0 \leq r^2 \leq 1$. El coeficiente de determinación multiplicado por 100 se puede interpretar como el porcentaje de pares de puntos que se pueden explicar con la recta de regresión.

Análisis de la Covarianza

El método estadístico que permite incluir la variable Grupo, que tiene escala nominal, es el análisis de la covarianza. El coeficiente estimado de la variable en escala nominal es el más importante en este análisis, por otro lado la variable en escala intervalar se denomina confundente.

El resultado del análisis de la covarianza arroja dos rectas de regresión, correspondientes a cada grupo (controles y pacientes), las dos rectas tienen la misma pendiente, y si la diferencia entre las rectas es estadísticamente significativa distinta de cero podemos decir que existe una diferencia entre los dos grupos.

Regresión múltiple

En los casos de más de una variable independiente la regresión se convierte en múltiple, con este método también es posible estudiar la interacción de dos variables calculando un coeficiente extra. El coeficiente de determinación múltiple cuantifica la cercanía de los puntos al plano de regresión; por otro lado el coeficiente de determinación ajustado tiene en cuenta la cantidad de coeficientes de la regresión múltiple, por lo tanto su valor es menor al r^2 múltiple.

Ambos coeficientes de determinación, junto con el estadístico F y el P asociado nos indican la validez de la regresión. En la mayoría de los casos de regresión múltiple no es posible una representación gráfica, en algunos casos se puede graficar en forma parcial las asociaciones.

Introducción al Método de Monte Carlo

Introducción

El método de Monte Carlo es una clase de métodos y algoritmos que en el caso de uso de modelización comienza con la generación de un conjunto de señales o secuencia de valores que se utilizan como variables de entrada del modelo, las cuales tienen propiedades pseudo aleatorias o estocásticas.

Estas variables son modeladas por generadores pseudoaleatorios, los cuales son algoritmos que proveen series determinísticas similares a series estocásticas.

La definición de un conjunto de números puramente aleatorios dice que dicho conjunto no debe tener correlación alguna con el modelo que luego lo utilizará como entrada, de forma que dos conjuntos aleatorios distintos deberían producir los mismos resultados cuando se aplican al mismo modelo. En la práctica se logran generadores pseudoaleatorios, los cuales tienen propiedades similares a un conjunto puramente aleatorio, tales como desviación uniforme.

Pruebas de aleatoriedad

- **Calculando los histogramas de cada serie pseudoaleatoria:** permite verificar gráficamente que tan uniforme es la distribución de los números dentro del rango determinado. Como principal limitación presenta la falta de un resultado numérico que indique la magnitud de la aleatoriedad.

- **Uso de la función autocorrelación (FAC):** permite calcular la correlación del conjunto de datos con sí mismo, mientras desplaza una copia del conjunto sobre la otra.

$$r_{xx}(j) = \frac{1}{N} \sum_{n=0}^{N-1} x(n)x(n+j)$$

donde r_{xx} es la función autocorrelación en el dominio del tiempo, de la secuencia x para el desplazamiento j , y N es la cantidad total de números de dicha secuencia. En el caso de utilizar más de un conjunto de datos aleatorios, se considera suficiente con calcular la FAC para cada una de ellas, pero cuando por ejemplo se utilizan dos conjuntos de datos pseudoaleatorios se puede calcular adicionalmente la correlación cruzada entre ambos (FCC)

$$r_{xy}(j) = \frac{1}{N} \sum_{n=0}^{N-1} x(n)y(n+j)$$

El resultado de la aplicación de la FAC sobre una señal pseudoaleatoria debería mostrar una correlación máxima (igual a uno), para un desplazamiento cero, y valores muy bajos de correlación para cualquier otro desplazamiento distinto de cero. Si la serie fuera puramente aleatoria obtendremos una correlación de uno sin desplazamiento y de cero para cualquier desplazamiento distinto de cero. Si la gran mayoría de las correlaciones para desplazamientos distintos de cero se encuentran dentro de los límites de significancia, el conjunto de datos se considera pseudoaleatorio.

$$Nivel_{significancia} = \pm \frac{z_{(1-\alpha)/2}}{\sqrt{N}}$$

Cálculo del área de un círculo con el método de Monte Carlo

Una forma simple de calcular el área del círculo de radio 1 con el método de Monte Carlo, se puede realizar encerrando dicho círculo dentro de un cuadrado con lados iguales a 2, entonces el área del cuadrado es por supuesto igual a 4. Luego aplicando el método de Monte Carlo podemos generar dos series pseudoaleatorias con una cantidad de números N , por ejemplo x e y , correspondientes a las coordenadas respectivas del gráfico del cuadrado y el círculo. Entonces si cada par de valores provenientes de los conjuntos x e y , representan las coordenadas de un punto dentro del cuadrado, podemos contar los puntos que caen dentro del círculo.

$$A_N^* = 4 \frac{n_{internos}}{N}$$

donde A_N^* es el área estimada para una cantidad de pares de datos N , 4 es el área del cuadrado, y $n_{internos}$ son la cantidad de puntos correspondientes al interior del círculo.

Resultados

Ahora vamos a aplicar el método de Monte Carlo para estimar el área de un círculo, para lo cual utilizaremos conjuntos de números con N desde 20 a 700, de forma de ver el efecto en el error de la estimación debido a la cantidad de números pseudoaleatorios.

Como podemos apreciar en el cuadro el error porcentual llega al 1 % con $N = 500$, e luego baja a un error menor al 0,15 % con $N = 700$.

N	20	50	100	250	500	700
A_N^*	3.4	3.2	3.32	3.264	3.176	3.137143
$e_N \%$	8.225	1.859	5.679	3.896	1.095	-0.142

En todos los casos la mayoría de las correlaciones se encuentran dentro de los niveles de significancia correspondiente para cada N , con lo cual se verifica la pseudoaleatoriedad de las secuencias.

Los valores de significancia para $\alpha = 0,05$ y dos colas, fueron de $\pm 0,438$, $\pm 0,277$, $\pm 0,196$, $\pm 0,124$, $\pm 0,0877$ y $\pm 0,074$, para los $N = 20, 50, 100, 250, 500$ y 700 , respectivamente.

Además como lo muestra el cuadro anterior la estimación mejora aumentando la cantidad N , y las figuras muestran que también disminuye la correlación y por lo tanto los conjuntos de números se pueden considerar más pseudoaleatorios con el aumento en el N .

Sistemas lineales e invariantes en el tiempo

Un sistema es lineal si obedece el principio de superposición:

- Si la entrada $x_1[n]$ produce la respuesta $y_1[n]$ y la entrada $x_2[n]$ produce la respuesta $y_2[n]$ con $n \in \mathbb{Z}$, \rightarrow la entrada ponderada y sumada $a_1x_1[n] + a_2x_2[n]$ produce la respuesta ponderada y sumada $a_1y_1[n] + a_2y_2[n]$ donde $a_1, a_2 \in \mathbb{R}$.

Un sistema es invariante en el tiempo si:

- la entrada $x[n]$ produce la respuesta $y[n]$ entonces la salida dada la entrada $x[n - N]$ produce la respuesta $y[n - N]$.
- La invarianza se refiere a todo desplazamiento de la función, aunque no estemos trabajando sobre el tiempo.

La convolución es conmutativa. La salida del sistema en función de una entrada cualquiera puede calcularse realizando la operación de convolución entre la entrada y la respuesta del sistema al impulso unitario.

Para señales discretas finitas (como los vectores que podemos almacenar en nuestras computadoras) debemos especificar cómo se maneja la convolución en los bordes del vector. La solución más común es considerar que cada vector es un período de una señal infinita y periódica. Entonces, cuando la operación de convolución accede a un elemento una posición más allá del final del vector, se accede al primer elemento del mismo (llamada convolución circular). Hay otras formas de lidiar con este problema, como agregar ceros antes y después del vector o reflejar el vector alrededor de sus bordes.

Convolución

Sistema

Cualquier proceso que produce una señal de salida en respuesta a una señal de entrada.

Homogeneidad

Un cambio de amplitud en la entrada produce un cambio idéntico de amplitud a la salida. Si $y[n] = S(x[n]) \rightarrow ky[n] = S(kx[n])$

Aditividad

Señales sumadas a la entrada producen señales sumadas a la salida.

Invarianza ante desplazamiento

Las características del sistema no cambian con el tiempo. Si $y[n] = S(x[n]) \rightarrow y[n + s] = S(x[n + s])$

LT: Sistema lineal e invariante ante desplazamiento.

Respuesta al impulso

Cualquier señal digital puede ser descompuesta en una serie de impulsos escalados y desplazados. Conocemos la salida o respuesta de un sistema a un impulso unitario en la entrada. Por lo tanto, y por la propiedad de superposición de los sistemas LTI, podemos conocer la respuesta de un sistema a cualquier señal de entrada a partir de su respuesta a un impulso unitario.

Convolución

Operación por la que obtenemos la señal de salida de un sistema a partir de la señal de entrada y la respuesta del sistema al impulso unitario.

Punto de vista de la entrada

Cada elemento de la señal de entrada contribuye a la señal de salida con una versión escalada y desplazada de la respuesta al impulso unitario.

Punto de vista de la salida

Cada elemento de la señal de salida recibe la contribución de varios elementos de la señal de entrada, multiplicados por la respuesta al impulso invertida.

Definición de convolución

$$y[n] = x[n] * h[n] = \sum_{k=0}^{N-1} x[k]h[n - k]$$

Correlación cruzada

Medida de similaridad de dos formas de onda, en función de un escalamiento, desplazamiento y suma de ruido blanco aplicados a una de ellas.

$$y[n] = \sum_{k=0}^{N-1} x[n]t[n + k]$$

Autocorrelación

Medida de similaridad entre “observaciones” en una misma señal, en función de la separación temporal entre ellas. Matemáticamente es la correlación cruzada de una señal con ella misma y es una herramienta para encontrar patrones que se repiten en una señal.

Filtro

Definición

Dispositivo o proceso que elimina un componente o una característica no deseada de una señal.

Usos

- Separar señales.
- Restaurar señales.

Tipos de filtro

1. Clasificación por respuesta
 - a. **FIR respuesta finita al impulso:** Estables. Los coeficientes b_k coinciden con los valores muestreados de la respuesta al impulso del filtro. El efecto de una entrada se extingue en la salida luego de un número finito de muestras. Por lo general implementados por convolución.
 - b. **IIR respuesta infinita al impulso:** La realimentación de la salida hace posible que la respuesta se extienda en forma ilimitada. Suelen ser más eficientes en cuanto a

consumo de recursos de cómputo y memoria que los filtros FIR. Pueden tener problemas de estabilidad y de propagación de errores numéricos.

2. Clasificación por filtrado de frecuencias

- a. Pasa bajos.
 - i. Un ejemplo es el FIR
 - ii. Unipolar pasa bajos (IIR):
$$y[n] = b_0.x[n] + a_1.y[n - 1]$$
 con $b_0 + a_1 = 1$
- b. Pasa altos.
 - i. Un ejemplo puede ser tomar la respuesta al impulso de un filtro pasa bajos con M impar (FIR), multiplicar su máscara por -1 y sumarle 1 a su elemento central.
 - ii. Unipolar pasa altos (IIR): $y[n] = b_0.x[n] + b_1.x[n - 1] + a_1.y[n - 1]$ con $b_0 = (1+a_1)/2$ y $b_1 = -b_0$
- c. Pasa banda.
- d. Rechaza banda.

Introducción a las Redes Neuronales: Procesamiento Adaptativo

Filtrado adaptativo

Son aquellos filtros autoprogramables cuya respuesta en frecuencia o función transferencia está alterada o adaptada para:

- dejar pasar sin degradación las componentes deseadas
- atenuar las señales no deseadas que generan interferencias
- reducir cualquier distorsión de señal de entrada.

Como generalmente no existe información disponible a priori, los filtros adaptativos requieren un período inicial de aprendizaje.

Estructura general de los filtros adaptativos

Compuesto por:

- El índice de performance, el cual debe ser optimizado
 - El índice de performance depende de la aplicación, es decir, está determinado por el tipo de trabajo del sistema adaptativo. En general se toma como evaluador la minimización del error cuadrático medio (ECM)
- El algoritmo que recalcula los parámetros del filtro: El algoritmo de optimización es el mecanismo mediante el cual los parámetros, optimizando el criterio, son calculados. Existen 2 tipos:
 - **No recursivo:** requiere la colección de todos los datos en una ventana temporal, para luego resolver un sistema de ecuaciones.
 - **Algoritmo del gradiente:** requiere la solución de un conjunto de ecuaciones lineales por inversión de una matriz y de esta forma los resultados no estarán disponibles en tiempo real.

Los algoritmos más utilizados desde el punto de vista del procesamiento digital son del tipo recursivo, el cual se actualiza a partir de cada entrada de la señal de entrada o con un pequeño grupo de muestras. Los resultados están disponibles inmediatamente y es posible el seguimiento de señales no estacionarias. Un ejemplo de ello es el algoritmo LMS.

- La estructura del filtro, que realiza las operaciones requeridas sobre la señal.
 - Depende tanto del algoritmo como de la aplicación.

Identificación de los parámetros

Para controlar el comportamiento óptimo de un sistema es necesario conocer su conducta dinámica. Esta está usualmente dada por ecuaciones diferenciales que relacionan la entrada y la salida del

sistema. Si sólo se reconoce la estructura de las ecuaciones, pero no los parámetros, algún tipo de algoritmo podría ser aplicado para estimarlos.

Las entradas y salidas ruidosas son medidas tanto durante la aplicación normal o durante la prueba de especificaciones. Estas son ingresadas a un filtro con coeficientes variables, los cuales están ajustados por un algoritmo optimizador. Luego de la adaptación, el filtro representa el mejor modelo para el sistema, y sus coeficientes son los parámetros identificados del sistema.

Corrección adaptativa de señales

El filtro de predicción lineal, es aquel que provee a su salida las salidas futuras del proceso. La misión específica de un filtro de predicción lineal es la de estimar la salida futura del proceso como una combinación lineal de las salidas presente y pasada.

La señal de entrada s , y versión retardada de la misma son enviadas al procesador adaptativo, el cual debe intentar predecir la entrada para hacer que con y y d se calcule el error, buscando que este último se minimice. Los filtros de predicción se utilizan además en decodificación de señales y reducción de ruido.

Técnicas de estimación óptima

El error es la resta (según la aplicación) entre la salida del sistema y la del modelo:

$$e = y - y_{\text{modelo}}$$

si para simplificar la nomenclatura hacemos $y_{\text{modelo}} = w$:

$$e_n = y_n - w_n$$

si tomamos para identificar un modelo estadístico tipo ARMA:

$$w_n = \sum_{i=1}^p \alpha'_i w_{n-i} + \sum_{j=0}^m \beta'_j x_{n-j}$$

Entonces para el error entre la salida del sistema físico y modelo se utiliza:

$$y_n = \sum_{i=1}^p \alpha_i y_{n-i} + \sum_{j=0}^m \beta_j x_{n-j}$$

$$e_n = \sum_{i=1}^p (\alpha_i - \alpha'_i) y_{n-i} + \sum_{j=0}^m (\beta_j - \beta'_j) x_{n-j}$$

El error cuadrático medio (ECM) es una magnitud que se calcula como:

$$\xi = E(e_n^2)$$

ξ posee un mínimo absoluto en el punto dado por:

$$\begin{aligned} \alpha_j &= \alpha'_j & 1 \leq i \leq p \\ \beta_j &= \beta'_j & 0 \leq j \leq m \end{aligned}$$

El problema de estimar los coeficientes se reduce a calcular los coeficientes α_i y β_j que minimizan el error cuadrático ξ

Método de resolución directa

Consisten en calcular el punto en que se anulan todas las derivadas del error cuadrático medio. Dado que la función ξ no posee máximos para los valores finitos de α_i y β_j la solución de las ecuaciones conducen al mínimo deseado.

Si se trata de un proceso cuasiestacionario, debe emplearse la esperanza matemática de ECM como la suma temporal dentro del período de estacionalidad del proceso, luego el cálculo debe repetirse periódicamente, de manera de cubrir todo el tiempo en estudio.

Métodos iterativos

Involucran un menor número de operaciones matemáticas y proveen además una estimación de α_i y β_j adaptable a las fluctuaciones de los coeficientes verdaderos α_i y β_j .

1. **Algoritmo de gradiente mediante el método de Newton:**

Consiste en descender sobre el hiperparaboloide de ECM siguiendo la dirección impuesta por el método de Newton-Raphson para la determinación de una raíz y por ende con una sencilla transformación, el mínimo.

$$f'(x_0) = \frac{f(x_0)}{x_0 - x_1} \quad \Rightarrow \quad x^{k+1} = x^k - \frac{f(x^k)}{f'(x^k)} \quad k = 0, 1, \dots$$

$$f(x) = \xi'(x)$$

\Rightarrow La convergencia depende del valor inicial y de la naturaleza de $f(x) \Rightarrow$

$$\alpha_i^{k+1} = \alpha_i^k - \mu \frac{\xi'(x)}{\xi''(x)} \quad 1 \leq i \leq p$$

$$\beta_j^{k+1} = \beta_j^k - \mu \frac{\xi'(x)}{\xi''(x)} \quad 0 \leq j \leq m$$

\Rightarrow en vez de un cero se tendrá un mínimo \Rightarrow

El coeficiente μ es una constante que determina la dimensión del peso, y gobierna la velocidad de convergencia, es decir un valor de μ demasiado pequeño requiere un elevado número de iteraciones para alcanzar el mínimo.

2. **Método de gradiente buscado por descenso escalonado**

En este caso los coeficientes o pesos son ajustados según el gradiente en cada paso. Este método está gobernado por las ecuaciones en la cual μ es una constante que regula el tamaño del paso. Tanto en este caso como en el anterior se parte utilizando todo el conocimiento previo sobre los valores a estimar y luego es desplazada la estimación inicial descendiendo sobre la superficie de ECM en la dirección del gradiente (perpendicular a las curvas de nivel) hasta alcanzar un mínimo. La diferencia entre este método y el de Newton-Raphson radica en la practicidad del algoritmo puesto que en el primero es necesario resolver una matriz inversa.

3. **Algoritmo LMS de Widrow**

En el algoritmo anterior se requiere la determinación del gradiente del ECM en cada iteración, es decir se estimó el gradiente de ξ . Bernard Widrow analizó las condiciones de estabilidad, la constante de tiempo de convergencia, y el error de estimación del algoritmo anterior, reemplazando el valor del gradiente por una estimación del mismo empleando una sola muestra. En otras palabras se toma a e_k^2 como una estimación directa de ξ

Aplicaciones del filtrado adaptativo

1. Modelización adaptativa en la síntesis de filtros FIR

La idea básica es asociar a un pseudofiltro las especificaciones ideales de un filtro que generalmente no será físicamente realizable.

a. Cancelación adaptativa de ruido

Una señal es transmitida de un canal a un sensor que recibe la señal más ruido no correlacionado n_0 . Esta combinación aditiva forma la primera entrada. Un segundo sensor recibe ruido n_1 el cual no está correlacionado con la señal pero sí lo está de

alguna manera con el ruido n_0 . Esta constituye la entrada de referencia al cancelador. El ruido n_1 es filtrado de forma tal de obtener una réplica n_0 . La salida se resta de la entrada primaria para producir una salida del sistema igual a $s + n_0 + y$.

2. Implementación de filtros adaptativos

a. Estructura básica de un filtro adaptativo

Los filtros adaptativos se pueden implementar tanto como FIR o IIR, pero en general se prefiere el esquema FIR debido a una mayor simplicidad y estabilidad, teniendo en cuenta que el algoritmo de adaptación es el encargado de calcular los coeficientes del filtro, por lo cual un esquema FIR asegura la estabilidad del filtro a pesar de las combinaciones de coeficientes calculadas. La estructura de un filtro FIR se puede desarrollar con un combinador lineal:

i. **Paralelo**

- ii. **Serie:** en la mayoría de las aplicaciones es necesario ingresar los datos (muestras) en forma serie, especialmente cuando el filtro se implementa en tiempo real y por lo tanto las muestras van llegando una a una.

El vector de entrada X se forma con la muestra actual $X(n)$ y las muestras anteriores $X(n - 1)$ a $X(n - 3)$, para lograr el retardo entre dichas muestras se intercala un bloque de retardo T .

El error cuadrático medio ξ describe una superficie de performance, cuando se lo grafica con respecto a los pesos W . El proceso de adaptación busca minimizar el error cuadrático medio, para lograr este objetivo se propone el método de descenso a pasos a través de la superficie de performance.

Se utiliza el error instantáneo porque está disponible muestra a muestra durante la optimización de los pesos del filtro.

3. Ejemplos de aplicaciones del filtrado adaptativo

a. Ejemplo de identificación de un sistema

Los filtros adaptativos son capaces de identificar un sistema desconocido, en otras palabras de desarrollar una modelización. Los pasos a seguir para lograr dicho proceso son:

- i. Ambos sistemas, el filtro y el sistema desconocido, deben ser excitados por un generador de ruido aleatorio de valor medio nulo.
- ii. El filtro adaptativo comienza un proceso de optimización, comenzando los pesos con valores iniciales aleatorios pequeños; utilizando el método de los cuadrados mínimos se optimizan los pesos (coeficientes del filtro) minimizando el error.
- iii. La cantidad de iteraciones necesarias para la optimización se pueden calcular buscando un error mínimo y una cantidad máxima de iteraciones, en caso de no lograr el error mínimo.

b. Ejemplo de cancelación de interferencias

En muchas ocasiones una señal de gran ancho de banda $w(n)$ se encuentra interferida por la suma de ruido de ancho de banda estrecho $s(n)$, conformando la señal suma $x(n)$; las dos señales que forman $x(n)$ no están correlacionadas. La solución a este problema es diseñar un filtro de ancho de banda estrecho (notch), pero en ciertas ocasiones la banda de ruido no es conocida o varía lentamente en el tiempo, para lo cual los filtros adaptativos pueden utilizarse, pues en cualquiera de las dos situaciones son capaces de optimizar su funcionamiento. El filtro adaptativo debe ser capaz de sustraer $s(n)$ de $x(n)$, teniendo en cuenta que debe predecir la componente en cada muestra a través de las muestras anteriores, dicha predicción se logra gracias al ancho de banda estrecho de $s(n)$, y debido a que como el ancho de banda de $s(n)$ es mucho menor que el de $w(n)$, las muestras de $s(n)$ tienen una alta correlación; por otro lado las muestras de $w(n)$ tienen una muy baja correlación.

La señal de error $e(n)$ se utiliza en el proceso de optimización del filtro adaptativo, con el método de los cuadrados mínimos. En dicho proceso de optimización $x(n)$ es la suma de una secuencia aleatoria de valor medio nulo, sobre la cual se suma el ruido de ancho de banda estrecha, por ejemplo un tono senoidal, o bien una suma de dos o tres tonos puros. El retardo del decorrelacionador en la mayoría de los casos es suficiente con una muestra de retardo. Una vez adaptado el filtro, la señal a filtrar debe ser ingresada directamente al filtro; eventualmente el filtro, puede reoptimizarse para seguir las variaciones de señal interferente.

Introducción a las Redes Neuronales: Red de Perceptrones

Introducción

Las RNA se basan en la interconexión de perceptrones, son máquinas o software que modelan la forma en la cual el cerebro desarrolla una tarea. Son una clase de métodos estadísticos para el procesamiento de información, compuestas por una cantidad importante de unidades simples de procesamiento, denominadas neuronas, las cuales interactúan entre sí.

Correspondencia entre la terminología de las RNA y la estadística:

RNA	Estadística
Nodos de entrada	Variables independientes
Nodos de salida	Variables dependientes
Aprendizaje	Parametrización
Pesos	Parámetros

Las dendritas son las entradas de información, y el axón es una salida mientras que mielinizado significa aislado.

Los atributos más importantes de las redes neuronales es el aprendizaje y la generalización. El aprendizaje es el proceso por el cual se le presenta a la red neuronal uno o varios patrones con sus salidas deseadas (tantos como sea posible), determinando un conjunto de entrenamiento, con dicho conjunto y mediante un algoritmo de optimización se ajustan los coeficientes de la red. Una vez entrenada la RNA, se puede probar la respuesta de la misma aplicando patrones distintos a los del entrenamiento, verificando si la respuesta fue correcta en la mayoría de los casos, entonces se puede decir que se obtuvo un entrenamiento general.

El perceptrón a partir de un combinador lineal adaptativo

El sistema completo, formado por un combinador lineal adaptativo y un elemento alineal a la salida de denomina adaline y tiene la propiedad de separabilidad lineal.

En una implementación práctica de una neurona, es conveniente incluir todas las alternativas posibles de cuantificación y cálculo error, de manera de poder implementar más de un algoritmo de optimización.

La elección por una u otra función de cuantificación depende de muchos factores, entre ellos cómo se implementan dichos cuantificadores. Por ejemplo si se construye una red neuronal con hardware, es mucho más fácil implementar la función signo.

Otra alternativa es utilizar la función sigmoidea durante el entrenamiento, porque con ella es posible llegar a mejores optimizaciones, para luego durante la operación de la red utilizar la función signo. Las redes neuronales se organizan en capas, formadas por una cantidad de neuronas cada una de ellas, generalmente son de dos o tres capas, denominándose las capas en las cuales entra el vector de entrada y donde se genera el vector de salida, capas de entrada y salida, respectivamente; las capas restantes entre la entrada y la salida se denominan capas ocultas.

Para el aprendizaje de una red neuronal, al igual que para una neurona simple, se puede utilizar el método de cuadrados mínimos, pero es necesario tener en cuenta para la optimización de todos los

pesos de todas las neuronas que conforman la red, el error producido a una salida deseada en el vector de salida.

El método de retropropagación busca propagar el error al valor deseado del vector de salida, en sentido inverso al flujo de información a través de la red neuronal; para lograr esto el error llega a las capas anteriores multiplicado por el peso de las capas posteriores.

Antes de comenzar el proceso de aprendizaje, generalmente se inicializan los pesos de todas las neuronas con valores aleatorios pequeños, después comienza el aprendizaje, el cual puede durar una cantidad de iteraciones determinadas previamente, o bien buscar un error de salida mínimo, o ambos métodos en caso de no lograr el error mínimo.

Transformada Discreta de Fourier

Cualquier señal continua y periódica puede ser representada como la suma de un conjunto de señales senoidales apropiadamente elegidas.

Las señales pueden ser clasificadas en:

- Continuas o discretas
- Periódicas o aperiódicas

La transformada de Fourier se puede clasificar en:

- **Transformada de Fourier:** señales periódicas y continuas.
- **Series de Fourier:** señales periódicas y continuas.
- **Transformada en tiempo discreto de Fourier:** señales aperiódicas y discretas.
- **Transformada discreta de Fourier:** señales periódicas y discretas.

Transformada real y compleja

Hay 2 tipos de Transformada real

- Directa: a partir de una señal $f : Z \rightarrow R$ obtenemos una señal $F : Z \rightarrow C$.
- Inversa: a partir de una señal $F : Z \rightarrow C$ obtenemos una señal $f : Z \rightarrow R$.

Hay 2 tipos de Transformada compleja

- Directa: a partir de una señal $f : Z \rightarrow C$ obtenemos una señal $F : Z \rightarrow C$.
- Inversa: a partir de una señal $F : Z \rightarrow C$ obtenemos una señal $f : Z \rightarrow C$.

El dominio de la frecuencia

- **Longitud de la señal:** Si la señal en el dominio del tiempo tiene N puntos (muestras) de longitud, en el dominio de la frecuencia tiene $N/2 + 1$ puntos de longitud.
- Se puede pensar la parte real de la transformada como un valor proporcional a la amplitud de las señales componentes coseno.
- Se puede pensar la parte imaginaria de la transformada como un valor proporcional a la amplitud de las señales componentes seno.
- Cada punto del dominio de la frecuencia corresponde a una señal senoidal de determinada frecuencia en el dominio del tiempo.

Variable independiente del dominio de la frecuencia

- Número de muestra: $ck[n] = \cos(2\pi kn/N)$
- Fracción de frecuencia de muestreo: $ck[n] = \cos(2\pi fn)$ con $f = k/N$
- Frecuencia natural: $ck[n] = \cos(\omega n)$ con $\omega = 2\pi f$ y $f_s \geq 2f_{max}$
- Frecuencia analógica particular a la aplicación (Hz).

Funciones base

- Conjunto de funciones seno y coseno de amplitud unitaria.
- Ecuaciones:

$$c_k[n] = \cos(2\pi \frac{kn}{N})$$

$$s_k[n] = \sin(2\pi \frac{kn}{N})$$

○

$$k = 0, 1..N/2 \text{ y } n = 0, 1..N - 1$$

○

- El parámetro k determina la frecuencia de cada senoide.
- $c_1[n]$ es la senoide que completa un ciclo en N puntos
- $c_{N/2}[n]$ es la senoide que completa N/2 ciclos en N puntos.

Puntos sin información

Si transformamos una señal con N muestras obtenemos N + 2 señales base de salida, puesto que el dominio de la frecuencia tiene N/2 + 1 puntos complejos. ¿Esto significa que tenemos más información en el dominio de la frecuencia que en el dominio del espacio?

El caso $k = 0$, hace $s_0[n] = 0$ siempre. El caso $k = N/2$ hace a $s_{N/2}[n] = 0$ siempre, puesto que el muestreo coincide con la mitad del ciclo del seno, es decir, donde éste cruza al cero. Por lo tanto, podemos ver intuitivamente que el dominio de la frecuencia no agrega más información que la que tenemos en el dominio del tiempo.

Transformada de Fourier inversa

La secuencia $F[n]$ de N números complejos, se transforma en la secuencia $f[n]$ de N números reales por medio de esta ecuación:

$$f[n] = \sum_{k=0}^{N/2} a_k \operatorname{Re}(F[k]) \cos(2\pi kn/N) + \sum_{k=0}^{N/2} b_k \operatorname{Im}(F[k]) \sin(2\pi kn/N)$$

Normalización de amplitud

$$a_0 = a_{N/2} = \frac{1}{N}$$

$$a_k = \frac{2}{N}$$

$$b_k = -\frac{2}{N}$$

Transformada de Fourier directa

$$F[k] = \sum_{n=0}^{N-1} f[n] \cos(2\pi kn/N) - \sum_{n=0}^{N-1} f[n] j \sin(2\pi kn/N)$$

Representación polar

Desplazamiento y suma de cosenos: En representación polar la suma de N/2 + 1 senos y N/2 + 1 cosenos se transforma en una suma de N/2 + 1 cosenos con desplazamiento de fase. Esto tiene

como ventaja que facilita la comprensión de las características de la señal al ver el dominio expresado en magnitud y fase.

Teorema de convolución

Este teorema afirma que bajo ciertas condiciones la transformada de Fourier de una convolución es el producto punto a punto de transformadas de Fourier.

$$F(x * y) = F(x) F(y)$$

Una de las ventajas de la transformada de Fourier es que existe un algoritmo llamado Transformada Rápida de Fourier (FFT) que permite realizar el cómputo en $O(n \log n)$ operaciones en vez de en $O(n^2)$ operaciones. Una de las restricciones para utilizar este algoritmo, es que el largo de los vectores de entrada deben ser potencias de 2. Si los vectores no tienen un largo potencia de 2, pueden agregarse ceros hasta llegar a la potencia de 2 más próxima.

Propiedades Transformada directa de Fourier

- Linealidad:
 - $TDF(af[n] + bf[n]) = aTDF(f[n]) + bTDF(f[n])$
- Dualidad:
 - Convolución en el tiempo: Multiplicación en la frecuencia.
 - Convolución en la frecuencia: Multiplicación en el tiempo
- La transformada de un impulso produce una función constante.
- La transformada de una función constante produce un impulso.

Órdenes de complejidad

- TDF
 - El cálculo de la TDF, tal como muestra su definición, requiere, por cada elemento de la señal en frecuencia, recorrer en forma completa el vector de elementos de la señal en el tiempo. Esto da como resultado un orden de complejidad $O(N^2)$.
- TRF
 - La transformada rápida de Fourier es un algoritmo que permite obtener la TDF en un tiempo proporcional a $O(N \log_2(N))$.
- Algoritmo Radix-2 de Cooley-Tukey
 - El único requisito de esta implementación es que el número de puntos de la serie debe ser potencia de 2. Si no tenemos suficientes datos, deben agregarse ceros.

Reducción de ruido espectral

El ruido en una señal que deseamos analizar puede hacer que su espectro de frecuencia sea ininteligible. El promediado de segmentos es una técnica que, a cambio de una reducción en la resolución del espectro, disminuye el ruido en el mismo.

Promediado de segmentos

Dado un número grande de muestras, se debe dividir la señal en segmentos iguales. Cada segmento debe ser multiplicado por una ventana adecuada (por ejemplo Hamming), para luego ser transformado y promediado con la magnitud de la transformada de los restantes segmentos. De esta forma el ruido aleatorio se reduce en forma proporcional a la raíz cuadrada del número de segmentos.

Series temporales

Definición

Una serie temporal es la representación de una variable en función del tiempo, durante un lapso determinado. Se pueden modelizar matemáticamente, descomponer y extraer información, e interpolar y extrapolar valores.

Aplicación

El análisis y modelización de series temporales permite dilucidar los mecanismos que las generan, extraer las contribuciones de los diferentes agentes que actúan sobre ellas, y por otro lado estimar

futuros comportamientos de dichas variables, basados en actuaciones anteriores. Por ej estimar futuros comportamientos de los precios de acciones en un mercado de valores.

Métodos de análisis

1. **Métodos en el dominio del tiempo**

- Incluyen el cálculo de parámetros estadísticos, tales como la media y el desvío estándar, el rango y el histograma. Los métodos en el dominio del tiempo trabajan directamente sobre las muestras de la serie temporal.
- Los métodos de cálculo de parámetros estadísticos se utilizan para representar características de las series temporales, otros métodos basados en filtros permiten extraer información de las series temporales, tales como sus componentes principales, tendencias, entre otras.
- **Filtros de media móvil:** Los filtros FIR de media móvil (también conocidos como filtros promediadores) calculan para cada muestra el promedio de una cantidad determinada de muestras a su alrededor. Ej de filtro de 3 muestras:

$$MA3(k) = \frac{a(k-1) + a(k) + a(k+1)}{3}$$

La salida del filtro para cada muestra k , que va desde 1 hasta $N - 1$, se calcula como el valor medio o promedio entre el valor de la serie a filtrar A para dicha muestra k , la muestra anterior $k - 1$ y la posterior $k + 1$.

- **Filtro IIR:** Los filtros IIR de respuesta infinita al impulso pueden lograr una mayor capacidad de filtrado, utilizando una menor cantidad de términos con respecto a los filtros de media móvil.

$$IIR2(k) = Coef_1 a(k) + Coef_2 IIR2(k-1)$$

Donde $Coef_1$ y $Coef_2$ son los coeficientes de filtro, cuya condición de estabilidad es $Coef_1 + Coef_2 = 1$ (La suma debe ser 1). El $Coef_1$ multiplica a la muestra de la serie A en su muestra actual k , mientras que el $Coef_2$ multiplica la salida anterior del filtro, es decir para $k - 1$.

2. **Métodos en el dominio de la frecuencia.**

- Están basados en métodos de estimación espectral tales como la transformada de Fourier y la modelización ARMA.
- La serie temporal A , fue construida combinando señales de frecuencias determinadas, denominadas A_0, A_1, A_2 y n

Teorema de muestreo de Nyquist-Shannon

El teorema demuestra que la reconstrucción exacta de una señal periódica continua en banda base a partir de sus muestras, es matemáticamente posible si la señal está limitada en banda y la tasa de muestreo es superior al doble de su ancho de banda.

Si la frecuencia más alta contenida en una señal analógica $x_a(t)$ es $F_{max} = B$ y la señal se muestrea a una tasa $F_s > 2F_{max} \equiv 2B$, entonces $x_a(t)$ se puede recuperar totalmente a partir de sus muestras.

Frecuencia de muestreo

La **tasa o frecuencia de muestreo** es el número de muestras por unidad de tiempo que se toman de una señal continua para producir una señal discreta, durante el proceso necesario para convertirla de analógica en digital.

Según el teorema de muestreo de Nyquist-Shannon, para poder replicar con exactitud (es decir, siendo matemáticamente reversible en su totalidad) la forma de una onda es necesario que la frecuencia de muestreo sea superior al doble de la máxima frecuencia a muestrear.

Efecto *aliasing*

Si se utiliza una frecuencia menor a la establecida por el teorema de Nyquist, se produce una distorsión conocida como *aliasing*; algunos autores traducen este término como *solapamiento*. El *aliasing* impide recuperar correctamente la señal cuando las muestras de ésta se obtienen a intervalos de tiempo demasiado largos. La forma de la onda recuperada presenta pendientes muy abruptas.

Filtro *antialiasing*

Para eliminar el *aliasing*, los sistemas de digitalización incluyen filtros paso bajo, que eliminan todas las frecuencias que sobrepasan la frecuencia crítica (la que corresponde a la mitad de la frecuencia de muestreo elegida) en la señal de entrada. Es decir, todas las frecuencias que queden por encima de la frecuencia máxima a muestrear seleccionada, son eliminadas. El filtro paso bajo para este uso concreto recibe el nombre de *filtro antialiasing*. Sin embargo, abusar de los filtros *antialiasing*, puede producir el mismo efecto que se quiere evitar