

# Muestreo Estocástico

Santiago Hernandez

November 2020

## 1 Introducción

En este capítulo nos concentraremos en un método original que hemos aplicado en conjunto a los algoritmos de estimación que fueron(/serán?) utilizados en el Capítulo .... y que ayudó a disminuir ostensiblemente la complejidad computacional del problema en cuestión. En lo que sigue esbozaremos una justificación teórica de la propuesta para luego exhibir resultados de las mejoras obtenidas en cuanto a desempeño.

En particular, nos interesamos por el problema de reducir la cantidad de muestras que representan a una dada señal sin perder la capacidad de distinguirla de otra señal contra la cual se la compara. Una buena medida de que tan distinguibles son dos señales dadas  $g(t)$  e  $g(t)$  se puede construir en base a una noción geométrica, más específicamente, a qué tan “alineadas” se encuentran. Esta noción es de especial interés cuando podemos considerar que trabajamos con espacios vectoriales (e.g., cuando utilizamos técnicas de procesamiento lineales), y no es otra cosa que el producto interno entre  $x$  e  $y$  que, en general, se lo puede definir como

$$\langle f, g \rangle = \int_T f(t)g^*(t)dt \quad (1)$$

adonde la integración se lleva a cabo en un cierto intervalo  $T$ . Nótese que, en el contexto adecuado, si multiplicamos por  $T^{-1}$  a la Ec. 1 se la puede considerar un estimador de la correlación cruzada entre dos procesos  $f$  e  $g$ , en el sentido del promedio temporal que aproxima al valor esperado de un proceso estocástico. En adelante, nos atenderemos a la correlación como medida de distinguibilidad/separabilidad de las señales con las que trabajamos.

En lo que sigue vamos a restringir el análisis a señales centradas alrededor de una portadora de frecuencia  $\omega_0$  y que pueden considerarse de banda angosta, i.e., que su ancho de banda  $W \ll \omega_0$ . Luego, para señales de la forma  $x(t) = \text{Re}\{a(t)e^{j\omega_0 t}\}$ , podemos decir que la envolvente compleja  $a(t) = \rho(t)e^{j\phi(t)}$  es tal que verifica que  $\rho(t) \approx \text{cte.}$  para intervalos de tiempo un orden de magnitud mayores al período de la portadora.

Como ejemplo para motivar el resto de la discusión, consideremos dos señales  $x$  e  $y$  adonde, por ejemplo,  $x$  es una señal de referencia nominal contra la que se quiere comparar una señal  $y$  recibida. En particular,

$$\begin{aligned}
x(t) &= \text{Re}\{ae^{j\omega_0 t}\} = a \cos(\omega_0 t) \\
y(t) &= \text{Re}\{\rho(t)e^{j\omega_0 t + \phi(t)}\} \approx b \cos(\omega_0 t + \phi(t)).
\end{aligned}$$

La correlación cruzada entre  $x$  e  $y$  viene dada por

$$\frac{1}{T} \int_0^T ab \cos(\omega_0 t) \cos(\omega_0 t + \phi(t)) dt \quad (2)$$

adonde  $\phi(t)$  podría ser, por ejemplo, un desfase contante  $\phi_0$  o podría tener un comportamiento arbitrario como función del tiempo. Analicemos el caso en que

$$\phi(t) \approx \phi_0 + \delta\omega t, \quad (3)$$

con  $\phi_0$  y  $\delta\omega$  constantes. Nótese que para señales de banda angosta adonde  $W \ll \omega_0$  la aproximación de la Ec. 3 puede resultar bastante general para tiempos que incluso representen varios períodos de la portadora, y en cuyo caso también valdrá que  $\delta\omega \ll \omega_0$ . Para tener una intuición del problema que estamos planteando,  $\phi(t)$  como el de la Ec. 3 podría estar representando el corrimiento Doppler de una señal recibida respecto a los valores nominales de frecuencia de la portadora.

Sin pérdida de generalidad, consideraremos  $\phi_0 = 0$ , en cuyo caso, la Ec. 2 resulta ser

$$\frac{1}{2} [\text{sinc}(2(2f_0 + \delta f)T) + \text{sinc}(2\delta f T)] \approx \frac{1}{2} [\text{sinc}(4f_0 T) + \text{sinc}(2\delta f T)] \quad (4)$$

con  $f_0 = \omega_0/(2\pi)$  y  $\delta f = \delta\omega/(2\pi)$  y habiendo usado la suposición  $\delta f \ll f_0$  para la aproximación. Si disponemos de un tiempo de observación  $T$  lo suficientemente grande, la Ec. .. nos dará que la correlación cruzada tiende a cero y es este número el que representa la relación que nos interesa entre las dos señales. Si consideramos adicionalmente que  $|\text{sinc}(x)| \leq \pi^{-1} \frac{1}{|x|}$  podemos deducir que la condición para obtener la correlación cruzada entre  $x$  e  $y$  con un error de a lo sumo  $\epsilon$  debe ser

$$T > \frac{1}{2\pi\delta f\epsilon} \quad (5)$$

adonde nuevamente hemos supuesto  $\delta f \ll f_0$ .

Si ahora pasamos a considerar señales de tiempo discreto que representan a las señales de interés muestreadas a una cierta tasa de muestreo  $f_s = \frac{1}{T_s}$ , entonces podemos aproximar a la Ec. 2 haciendo

$$R_{xy} := T_s \sum_{n=0}^{(N-1)T_s} x(nT_s)y(nT_s) \quad (6)$$

con  $T = nT_s$ .

Estamos en condiciones de enunciar el problema que motiva el presente capítulo: *por un lado, deseamos que la Ec. 6 sea una buena aproximación de la Ec. 4 para un error  $\epsilon$  suficientemente pequeño, por otro, queremos que la cantidad de muestras  $N$  sea lo más chica posible para disminuir la carga computacional.*

Antes de enunciar la proposición que motiva el presente capítulo, enumeremos algunas alternativas *naive* para intentar dar una solución al problema planteado:

- Dado un cierto  $T_s$  fijo que hace que la Ec. 6 pueda considerarse una buena aproximación de Ec. 4, podríamos escoger un  $N_1 < N$ , adonde  $NT_s = T$ , pero en este caso no podemos cumplir con la condición de la Ec. 5. Más aún, de la forma misma que adopta la correlación cruzada, se puede ver que si achicamos mucho el número de muestras consideradas, se puede llegar a valores de correlación cruzada que indiquen una considerable correlación entre las señales comparadas, y esto es exactamente lo opuesto a lo que debe manifestarse.
- Si fijamos el intervalo de integración  $T$  garantizamos que se verifica la Ec. 5 en tiempo continuo, pero para pasar a discreto, si queremos achicar  $N$ , luego nos vemos obligados a aumentar  $T_s$ . De este modo la aproximación de Ec. 6 a la Ec. 4 es el factor limitante y con la disminución del número de muestras rápidamente se pierde convergencia al valor buscado. Para visualizar esto de un modo intuitivo pensemos lo siguiente: si bajamos la frecuencia de muestreo  $f_s$  al punto de dejar de cumplir con el teorema de muestreo de Nyquist, luego se manifestarán en el espectro de nuestra señal discretizada alias correspondientes a frecuencias menores. Como vimos de la Ec. 4 y la condición de la Ec. 5, cuanto menores sean las frecuencias involucradas, el tiempo  $T = NT_s$  necesario para poder distinguirlas (i.e., que la correlación cruzada converja a menos del error  $\epsilon$ ) será mayor.

Como hemos visto, disminuir la cantidad de muestras ya sea achicando el intervalo de integración numérica o, si dejamos este último fijo, aumentando el tiempo entre muestras, no conduce a los resultados esperados. Intuitivamente, si queremos achicar la cantidad de muestras pero conservando el intervalo total de observación y al mismo tiempo no perder información que ocurre a escala temporales pequeñas, no parece haber alternativa a utilizar algún tipo de distribución de tiempos entre muestras que no sea uniforme. En particular, como a priori se desea trabajar con señales genéricas, lo que se propone es que esta distribución siga algún tipo de ley aleatoria que no favorezca ningún escenario en particular. Esto nos lleva a enunciar el algoritmo de *muestreo aleatorio* que desarrollaremos en la próxima sección.

como un último remark, cabe destacar que al quitar muestras cosas que están correlacionadas siguen correlacionadas mientras que cosas que están decorrelacionadas no, es por esto que el análisis llevado a cabo se para en este eslabón débil para motivar el método.

## 2 Muestreo aleatorio

Teniendo en mente la intuición que surgió de la motivación de la sección anterior, consideraremos lo que sucede con una versión en tiempo discreto de la señal  $z(t) = x(t)y(t)$ . Consideraremos como señal muestreada a

$$z_m(t) = z(t)s_m(t) \quad (7)$$

adonde

$$s_m(t) = \left( \sum_{n \in \mathbb{Z}} \delta(t - nT_s) \right) m(t) \quad (8)$$

y  $m(t)$  es una *máscara de muestreo*. Si, por ejemplo,  $m(t) = 1$  para todo  $t$ , luego tendremos que  $z_m(t)$  se corresponde a  $z(t)$  muestreada idealmente y la correlación de la Ec. 6 se puede calcular como

$$\frac{T_s}{T} \int_0^T z_m(t) dt \quad (9)$$

de forma exacta. Sin embargo, como ya adelantamos, nos interesa utilizar una forma de muestreo aleatorio que consiste en definir la máscara  $m(t)$  como una realización del proceso estocástico definido por

$$M(t) = \sum_{n \in \mathbb{Z}} \Pi\left(\frac{t - nT_s}{T_s}\right) \cdot c_n \quad (10)$$

siendo

$$\Pi(t) = \begin{cases} 1 & t \in [0, 1) \\ 0 & \text{otro } t \end{cases}$$

y  $c_n$  un proceso aleatorio discreto tal que  $c_n = 1$  con probabilidad  $p$ ,  $c_n = 0$  con probabilidad  $1 - p$  y  $c_n$  es independiente de  $c_m$  para todo  $n \neq m$ . Nótese que para  $p = 1$  estamos en el caso de muestreo ideal y para  $p = 0$  no se utiliza ninguna muestra. El valor esperado de muestras utilizando este enfoque es de  $pT/T_s$ .

La clave está en notar que la correlación que buscamos aproximar (Ec. 9) se corresponde a calcular numéricamente el valor medio temporal de la señal  $z(t)$  a partir de un instante inicial  $t = 0$ . Es decir, nos interesa específicamente lo que ocurre con la componente de continua de dicha señal. Por otro lado, dada una realización  $m(t)$  del proceso  $M(t)$  para una cierta probabilidad  $p$ , podemos reescribir

$$m(t) = \mu_m + m_0(t) \quad (11)$$

con  $\mu_m$  una constante que representa el valor medio de  $m(t)$  y cuyo valor esperado es  $p$ , y  $m_0(t) = m(t) - \mu_m$ , siendo  $m_0(t)$  una realización del proceso  $M(t) - p$ , es decir, un proceso con las mismas características que  $M(t)$  pero de media nula.

Estamos en condiciones de calcular Ec. 9 que comprende el resultado principal de nuestra propuesta. Esto es

$$\frac{T_s}{T} \int_0^T z_m(t) dt = \mu_m R_{xy} + \frac{T_s}{T} \int_0^T z(t) m_0(t) dt \quad (12)$$

El primer término es la correlación que buscamos determinar a menos de un factor  $\mu_m$  conocido y el segundo término representa “ruido” que se agrega por la utilización del método de muestreo aleatorio propuesto. Para mayor claridad, si multiplicamos por  $\mu_m^{-1}$  a ambos lados de la igualdad, tenemos

$$\mu_m^{-1} \left( \frac{T_s}{T} \int_0^T z_m(t) dt \right) = R_{xy} + \mu_m^{-1} \left( \frac{T_s}{T} \int_0^T z(t) m_0(t) dt \right) \quad (13)$$

lo que deja en evidencia que al disminuir el valor de  $\mu_m$  (y, por lo tanto, utilizar menos muestras), también se amplifica el valor del ruido (recordemos que  $0 < \mu_m \leq 1$ ). También podemos estimar muy bien cuanto vale ese ruido en la determinación de  $R_{xy}$ . Suponiendo que  $T_s \ll T$  y que  $z(t)$  es independiente de  $m_0(t)$ , se puede mostrar fácilmente que

$$\left| \int_0^T z(t) m_0(t) dt \right| \leq \frac{M_z T_s}{2} \quad (14)$$

con  $M_z = \max_{0 \leq t \leq T} (|z(t)|)$ , de modo que el ruido en la estimación de la correlación  $R_{xy}$  por el método de muestreo aleatorio queda acotado por:

$$\mu_m^{-1} \frac{M_z T_s^2}{2 T}. \quad (15)$$

Considerando que  $N = \mu_m T / T_s$  es el número de muestras utilizando muestreo aleatorio, tenemos que el error en la estimación estará acotado por

$$\frac{M_z T_s}{2 N}. \quad (16)$$

A modo de ejemplo, en la Fig. 1 presentamos el cálculo numérico de la correlación  $R_{xy}$  para señales  $x$  e  $y$  como las presentadas en la sección anterior para  $\delta\omega = \omega_0/10$  y utilizando muestreos aleatorios con distintos valores de  $p$  (i.e.  $\approx \mu_m$ ). Las funciones de correlación están calculadas para  $T$  variando entre 0 y un valor suficientemente grande. El valor de correlación que nos interesa es para  $T$  grande. Como se puede ver, tanto para  $p$  grandes como pequeños, la aproximación es muy buena.

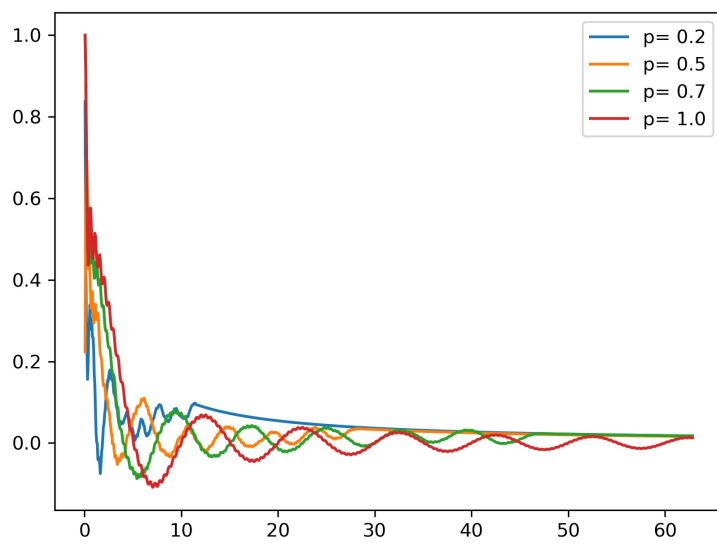


Figure 1: completar