Tarea 03 - Procesamiento Adaptativo

Resumen: Secciones 6.1, 6.2, 6.4 y 6.5 [1].

Fabricio Quirós Corella

30 de noviembre de 2019

Las aproximaciones estocásticas del gradiente en métodos de rápido descenso, empleados en implementaciones de filtros adaptativos, introducen ruido de gradiente, por lo que consecuentemente el rendimiento de dichos filtros se verá disminuido en comparación a los algoritmos originales de descenso rápido. El presente documento procura resumir un conjunto de criterios [1], basados en un análisis de estado estable, y utilizados para evaluar uniformemente el rendimiento de filtros adaptativos de distintos tipos.

Inicialmente, el estudio del rendimiento de un filtro adaptativo resulta necesario tratarlo como un proceso estotástico, en lugar de uno determinístico, lo que implica que las variables asociadas deben ser tomadas como aleatorias en la ecuación de recursión de mínimos cuadrados; es decir, un filtro LMS (*least mean-square*) con sus estimadores y observaciones definidos de forma aleatoria, los cuales están representados en negrita de la siguiente forma:

$$\mathbf{w_i} = \mathbf{w_{i-1}} + \mu \mathbf{u_i^*}[\mathbf{d}(i) - \mathbf{u_i w_{i-1}}] \tag{1}$$

El criterio más popular en cuánto a comparación de rendimiento de filtros adaptativo se refiere, corresponde al error cuadrático medio (en inglés, MSE) de estado estable, el cual está definido por la siguiente ecuación:

$$\text{MSE} \triangleq \lim_{i\to\infty} \text{E}|\mathbf{e}(i)|^2 \tag{2}$$

Donde, $e(i)$ corresponde al error de salida estimado *a priori*, equivalente a:

$$\mathbf{e}(i) \triangleq \mathbf{d}(i) - \mathbf{u_i w_{i-1}} \tag{3}$$

Y la función de costo mínimo del problema de mínimos cuadrados corresponde a la expresión:

$$J_{min} = \sigma_d^2 - R_{ud}R_u^{-1}R_{du} \tag{4}$$

Por lo que es común definir el *exceso del error de mínimos cuadrados* (en inglés, EMSE) de un filtro adaptativo como la diferencia:

$$\text{EMSE} \triangleq \text{MSE} - J_{min} \tag{5}$$

Además, una medida relativa del rendimiento de un filtro adaptativo denominada *desajuste* se define como:

$$\mathcal{M} \triangleq \frac{\text{EMSE}}{J_{min}} \tag{6}$$

Por consiguiente, resulta de interés evaluar el EMSE de un filtro adaptativo descrito por una ecuación de diferencias estocástica, esto adicionalmente mediante el modelado de los datos $\{\mathbf{d}(i), \mathbf{u_i}\}$.

Retomando el principio de ortogonalidad de la estimación lineal de mínimos cuadrados, es posible denotar el error de estimación resultante en función de la estimación óptima w^o, como $\mathbf{v}(i) = \mathbf{d}(i) - \mathbf{u_i}w^o$.

Ahora, es posible reescribir el resultado anterior mediante un modelo lineal para algún w^o que relaciona los datos $\{\mathbf{d}(i), \mathbf{u_i}\}$:

$$\mathbf{d}(i) = \mathbf{u_i}w^o + \mathbf{v}(i) \tag{7}$$

Como la señal $\mathbf{v}(i)$ no está correlacionada con $\mathbf{u_i}$, implica que la varianza de $\mathbf{v}(i)$ equivale a la función de mínimo costo J_{min}, así:

$$\sigma_v^2 \triangleq \mathrm{E}|\mathbf{v}(i)|^2 = J_{min} = \sigma_d^2 - R_{ud}R_u^{-1}R_{du} \tag{8}$$

Adicionalmente, con el propósito de hacer más manejable el análisis del rendimiento de los filtros adaptativos, se asume que los datos $\{\mathbf{d}(i), \mathbf{u_i}\}$ satisfacen las siguientes condiciones:

a. Existe un vector w^o que satisface la expersión $\mathbf{d}(i) = \mathbf{u_i}w^o + \mathbf{v}(i)$.

b. La secuencia de ruido $\{\mathbf{v}(i)\}$ es *independiente e identicamente distribuida* (i.i.d), con varianza $\sigma_v^2 = \mathrm{E}|\mathbf{v}(i)|^2$.

c. La secuencia $\{\mathbf{v}(i)\}$ es independiente de $\mathbf{u_j}$, para todo i, j.

d. La condición inicial $\mathbf{w}_{-1}$ es independiente de todo $\{\mathbf{d}(j), \mathbf{u_j}, \mathbf{v}(j)\}$.

e. La matriz de covarianza del regresor es $R_u = \mathrm{E}\mathbf{u_i}^*\mathbf{u_i} > 0$.

f. Las variables aleatorias $\{\mathbf{d}(i), \mathbf{u_i}, \mathbf{v}(i)\}$ poseen media igual a 0.

El modelo anterior constituye la descripción de un ambiente *estacionario*; es decir, uno con cantidades constantes para $\{w^o, R_u, \sigma_v^2\}$, donde una importante consecuencia de esto es el hecho que en cualquier instante de tiempo i, la variable del ruido $\mathbf{v}(i)$ será indepediente de todos los valores estimados previamente $\{\mathbf{w_j}, j < i\}$. Dado lo anterior, la secuencia $\mathbf{v}(i)$ de forma similar es independiente de $\{\widetilde{\mathbf{w}}_\mathbf{j}, j < i\}$, donde $\widetilde{\mathbf{w}}_\mathbf{j}$ denota al vector de error ponderado, definido así:

$$\widetilde{\mathbf{w}}_\mathbf{j} \triangleq w^o - \mathbf{w_j} \tag{9}$$

Además, el vector de ruido $\mathbf{v}(i)$ es también independiente de la estimación del error *a priori* $\mathbf{e_a}(i)$, dado por:

$$\mathbf{e_a}(i) \triangleq \mathbf{u_i}\widetilde{\mathbf{w}}_{\mathbf{i-1}} \tag{10}$$

Dicha variable mide la cercanía del estimador $\mathbf{u_i}\mathbf{w_{i-1}}$ al estimador lineal óptimo de $\mathbf{d}(i)$, denominado $\hat{\mathbf{d}}(i) = \mathbf{u_i}w^o$.

Considerando el modelo estacionario expuesto, es posible deducir una expresión más compacta para el EMSE de un filtro adaptativo. Primero, tomando en cuenta la Ec. 2, la Ec. 5 y la Ec. 8, se tiene por definición que:

$$\text{EMSE} = \lim_{i\to\infty} \text{E}|\mathbf{e}(i)|^2 - \sigma_v^2 \tag{11}$$

Ahora, usando la Ec. 3 y el modelo lineal de la Ec. 7, es posible encontrar que:

$$\mathbf{e}(i) = \mathbf{v}(i) + \mathbf{u_i}(w^o - \mathbf{w_{i-1}}) \tag{12}$$

Sustituyendo la Ec. 9 y la Ec. 10 en la expresión anterior:

$$\mathbf{e}(i) = \mathbf{v}(i) + \mathbf{u_i}\widetilde{\mathbf{w}}_{\mathbf{i-1}} = \mathbf{v}(i) + \mathbf{e_a}(i) \tag{13}$$

Tomando en cuenta la independencia de $\mathbf{v}(i)$ y $\mathbf{e_a}(i)$, se obtiene de la Ec. 13 que:

$$\text{E}|\mathbf{e}(i)|^2 = \text{E}|\mathbf{v}(i)|^2 + \text{E}|\mathbf{e_a}(i)|^2 = \sigma_v^2 + \text{E}|\mathbf{e_a}(i)|^2 \tag{14}$$

Así, evaluando en la Ec. 11, se deduce una expresión alternativa para el EMSE:

$$\text{EMSE} = \lim_{i\to\infty} \text{E}|\mathbf{e_a}(i)|^2 \tag{15}$$

Resulta evidente que el EMSE puede ser calculado al evaluar el valor medio cuadrático en estado estacionario de la estimación de error *a priori*. De la misma manera, a partir de la Ec. 14, es posible reescribir el MSE como:

$$\text{MSE} = \text{EMSE} + \sigma_v^2 \tag{16}$$

De la misma manera, basándose en el modelo lineal descrito, el cual permite identificar la función de costo mínimo J_{min} como σ_v^2, la razón de desajuste quedaría definida así:

$$\mathcal{M} = \frac{\text{EMSE}}{\sigma_v^2} \tag{17}$$

Por otro lado, en el estudio del rendimiento en estado estable, es necesario describir la operación de un filtro adaptativo en dicho estado estacionario, el cual debe cumplir las siguientes condiciones:

$$\text{E}\ \widetilde{\mathbf{w}}_{\mathbf{i}} = \text{E}\ \widetilde{\mathbf{w}}_{\mathbf{i-1}} = s, \text{ cuando } i \to \infty \text{ (usualmente } s = 0) \tag{18}$$

$$\text{E}\ \widetilde{\mathbf{w}}_{\mathbf{i}}\widetilde{\mathbf{w}}_{\mathbf{i}}^* = \text{E}\ \widetilde{\mathbf{w}}_{\mathbf{i-1}}\widetilde{\mathbf{w}}_{\mathbf{i-1}}^* = C, \text{ cuando } i \to \infty \tag{19}$$

Esto implica que la media y la matriz de covarianza del vector del error ponderado tienden a valores constante finitos. En particular, esto sugiere que se mantiene la siguiente condición:

$$\mathrm{E}\ ||\widetilde{\mathbf{w}}_{\mathbf{i}}||^2 = \mathrm{E}\ ||\widetilde{\mathbf{w}}_{\mathbf{i-1}}||^2 = c < \infty, \text{ cuando } i \to \infty \tag{20}$$

Donde, $c = \mathrm{Tr}(C)$.

Referencias

[1] Ali H Sayed. *Fundamentals of adaptive filtering.* John Wiley & Sons, 2003.