ADAPTİF FİLTRELERDE GAUSS-SEIDEL ALGORİTMASININ STOKASTİK YAKINSAMA ANALİZİ

Metin HATUN [*]
Osman Hilmi KOÇAL [*]

Özet: Bu makalede, adaptif filtre parametrelerinin iteratif olarak ayarlanmasında kullanmak için önerilen Gauss-Seidel algoritmasının deterministik ve stokastik yakınsama analizi yapılmıştır ve Normal denklemlerin optimum çözümünü veren yansız bir kestireç olduğu gösterilmiştir.

Anahtar Kelimeler: Adaptif Filtreler, Normal Denklem, Gauss-Seidel İterasyonu, Stokastik Algoritma, Yansız Kestireç, Yakınsama Analizi.

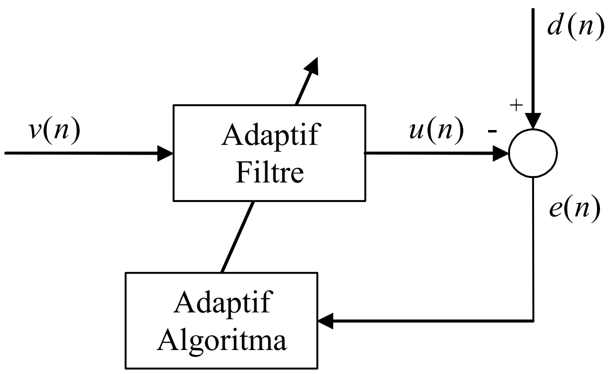
Stochastic Convergence Analysis of Gauss-Seidel Algorithm in Adaptive Filters

Abstract: In this article, deterministic and stochastic convergence analysis of Gauss-Seidel algorithm that is proposed for adjusting adaptive filter parameters in iterative manner is performed, and it is shown that it is an unbiased parameter estimator that gives the optimum solution for Normal equations.

Key Words: Adaptive Filters, Normal Equation, Gauss-Seidel Iteration, Stochastic Algorithm, Unbiased Estimator, Convergence Analysis.

1. GİRİŞ

Bilindiği gibi, adaptif filtre parametrelerinin iteratif olarak ayarlanmasında kullanılan algoritmalar optimizasyon tabanlı algoritmalar ve ardışıl en küçük kareler tabanlı algoritmalar olmak üzere kabaca iki gruba ayrılabilir (Haykin,1991, Farhang-Boroujeny, 1998). Bu algoritmalar aşağıdaki şekilde görüldüğü gibi filtre parametrelerini iteratif olarak ayarlamak için kullanılır.



Şekil 1:
Adaptif filtreleme işlemi

Burada $v(n)$ adaptif filtrenin ayrık-zamanlı giriş sinyali, $u(n)$ adaptif filtrenin tahmin edilen parametreler kullanılarak hesaplanan çıkış sinyali, $d(n)$ ise adaptif filtrenin takip etmesi istenen sinyaldir (Haykin, 1991, Farhang-Boroujeny, 1998). Parametre ayarlama algoritmalarıyla, tahmin edilen parametre-

[*] Uludağ Üniversitesi, Mühendislik-Mimarlık Fakültesi, Elektronik Mühendisliği Bölümü, Görükle, Bursa.

leri kullanarak hesaplanan $u(n)$ sinyali ile adaptif filtrenin izlemesi istenen $d(n)$ sinyali arasındaki hatanın karesinin beklenen değerini minimum yapan parametre tahminleri iteratif olarak hesaplanır (Haykin, 1991, Farhang-Boroujeny, 1998). Son yıllarda adaptif filtre katsayılarını iteratif olarak hesaplamak için, doğrusal denklem takımlarının iteratif çözümü üzerine kurulu bazı algoritmalar önerilmiştir (Koçal, 1998, Ng ve diğ., 2003, Albu ve Kwan, 2004, Bouchard ve Albu, 2005). Bunun için çoğunlukla Gauss-Seidel iterasyonu kullanılmaktadır. Önerilen bu algoritmalarda, Gauss-Seidel algoritması Normal denklemi iteratif olarak çözerek optimum parametre tahminlerini hesaplamak için kullanılmaktadır. Bu yöntemlerin avantajı, parametre güncelleme işleminin skaler olarak yapılmasıdır (Koçal, 1998, Ng ve diğ., 2003, Albu ve Kwan, 2004, Bouchard ve Albu, 2005). Bu makalelerde ve bu makalelerin referanslarında Gauss-Seidel algoritmasının stokastik yakınsama analizi yapılmamıştır. Bu makalenin amacı, adaptif filtre parametrelerini iteratif olarak hesaplamak için kullanılan bu algoritmaların deterministik ve özellikle stokastik yakınsama analizini yapmaktır.

2. GAUSS-SEIDEL ALGORİTMASI KULLANARAK PARAMETRE TAHMİNİ

Gauss-Seidel yöntemlerindeki orijinal fikir, en küçük kareler tahminlerinin elde edilmesinde, lineer denklem takımlarının çözümünde kullanılan iteratif yöntemlerden yararlanmaktır. Gauss-Seidel algoritması

$$R w_0 = p \tag{1}$$

olarak verilen normal denklemlerin çözümünde kullanılmaktadır. Burada, M adaptif filtrenin derecesi olmak üzere, R ile $M \times M$ boyutlu korelasyon matrisi, p ile $M \times 1$ boyutlu korelasyon vektörü, w_0 ile optimum parametre tahminleri gösterilmektedir. Buradaki korelasyon matrisi, korelasyon vektörü ve filtre çıkış verilerini içeren $\underline{u}(n)$ vektörü ise sırasıyla

$$R = E\left(\underline{u}(n)\underline{u}^T(n)\right) \tag{2}$$

$$p = E\left(\underline{u}(n)d(n)\right) \tag{3}$$

$$\underline{u}(n) = \begin{bmatrix} u(n-1) & u(n-2) & \cdots & u(n-M) \end{bmatrix}^T \tag{4}$$

olarak tanımlanmaktadır (Haykin, 1991). Burada $E(\cdot)$ beklenen değer operatörünü göstermektedir. Korelasyon matrisi ve korelasyon vektörü pratik uygulamalarda

$$R(N) \cong \frac{1}{N}\sum_{n=1}^{N}\underline{u}(n)\underline{u}^T(n) \tag{5}$$

$$p(N) \cong \frac{1}{N}\sum_{n=1}^{N}\underline{u}(n)d(n) \tag{6}$$

şeklinde N adımlık veri grubu kullanılarak tahmin edilmektedir veya

$$R(n) = R(n-1) + \underline{u}(n)\underline{u}^T(n) \tag{7}$$

$$p(n) = p(n) + \underline{u}(n)d(n) \tag{8}$$

şeklinde veya buna benzer şekillerde yeni örnekler kullanılarak iteratif olarak güncellenmektedir (Haykin, 1991). Gauss-Seidel iterasyonu bu aşamada iki farklı biçimde kullanılmaktadır. Birincisinde, korelasyon matrisi ve korelasyon vektörü N adet veri grubu kullanılarak tahmin edildikten sonra veya veriler her adımda elde edildikten sonra iteratif olarak güncellenmekte, daha sonra Gauss-Seidel iterasyonu veya bazı hızlandırılmış versiyonları kullanılarak Normal denklemi sağlayan optimum parametre tahminleri iteratif olarak hesaplanmaktadır (Ng ve diğ., 2003). İkincisinde ise, korelasyon matrisi ve korelasyon vektörü veriler her adımda elde edildikten sonra iteratif olarak güncellenmekte ve Gauss-Seidel algoritması tek başına veya bazı algoritmalarla birlikte kullanılarak adaptif filtre parametreleri güncellenmektedir (Koçal, 1998, Albu ve Kwan, 2004, Bouchard ve Albu, 2005).

3. GAUSS-SEIDEL ALGORİTMASI'NIN YAKINSAMA ANALİZİ

Gauss-Seidel algoritmasıyla parametre tahmin işleminde başlangıç noktası, zaman-ortalamalı Normal denklemlerin Gauss-Seidel algoritması ile çözümü üzerine kuruludur. Korelasyon matrisinin ve korelasyon vektörünün değerinin bilinmemesi durumunda, tahmini değeri kullanılarak bir yaklaşıklık yapılır. Korelasyon matrisi ve korelasyon vektörü eldeki veriler kullanılarak tahmin edildikten sonra veya elde edilen yeni veriler kullanılarak iteratif olarak güncellendikten sonra, Gauss-Seidel algoritması normal denklemlerin çözümünde kullanıldığında, parametre vektörü

$$w(n+1) = -(R_L + R_D)^{-1} R_U w(n) + (R_L + R_D)^{-1} p \tag{9}$$

iterasyonu kullanarak güncellenmektedir. Burada R_D korelasyon matrisinin köşegen elemanlarını içeren, R_L korelasyon matrisinin alt üçgen elemanlarını içeren, R_U ise korelasyon matrisinin üst üçgen elemanlarını içeren kare matrislerdir (Koçal, 1998). Bu aşamada Gauss-Seidel algoritmasının yakınsama analizi, deterministik yakınsama analizi ve stokastik yakınsama analizi olmak üzere iki farklı şekilde yapılabilir.

3.1. Deterministik Yakınsama Analizi

Korelasyon matrisinin ve korelasyon vektörünün (5) ve (6) eşitliklerindeki gibi N adet veri grubu kullanılarak tahmin dildiğini varsayalım. Bu durumda (9) eşitliğini

$$w(n+1) = -(R_L(N) + R_D(N))^{-1} R_U(N) w(n) + (R_L(N) + R_D(N))^{-1} p(N) \tag{10}$$

şeklinde yazabiliriz. Burada n iterasyon adımını göstermektedir. Bilindiği gibi R korelasyon matrisi simetrik ve pozitif tanımlı olan bir matristir (Haykin, 1991). Doğrusal denklem takımını oluşturan kare matrisin simetrik ve pozitif tanımlı olması durumunda, Gauss-Seidel algoritmasının herhangi bir başlangıç değeri için denklem takımını sağlayan çözüm değerine iteratif olarak yakınsadığı matematiksel olarak Golub ve Van Loan (1996) tarafından gösterilmiştir. Buradan hareketle Gauss-Seidel algoritmasının deterministik olarak normal denklemlerin optimum çözümüne yakınsadığı aşağıdaki lemma ile gösterilebilir:

Lemma: $R \in \Re^{M\times M}$ korelasyon matrisi simetrik ve pozitif tanımlı olduğundan dolayı, Gauss-Seidel iterasyonu herhangi bir $w(0)$ başlangıç değeri için normal denklemlerin w_0 ile verilen optimum çözümüne yakınsar.

İspat: Burada öncelikle yeterli miktarda veri kullanıldığı durumda

$$E(R(N)) \cong R \ , \ E(R_D(N)) \cong R_D \ , \ E(R_L(N)) \cong R_L \ , \ E(p(N)) \cong p \tag{11}$$

yaklaşıklıklarını yapabiliriz. R korelasyon matrisini

$$R = R_L + R_D + R_L^T \tag{12}$$

şeklinde yazdığımızda, (9) ile verilen iteratif algoritmadaki

$$T = -(R_D + R_L)^{-1} R_L^T \tag{13}$$

matrisinin bütün özdeğerleri birim daire içinde ise algoritma yakınsar. Bunun için spektral yarıçap olarak bilinen ve T matrisinin en büyük özdeğeri olarak tanımlanan $\max(|\lambda_T|) < 1$ olmalıdır. R_D matrisi pozitif tanımlı olduğundan dolayı

$$T_1 = R_D^{1/2}\, T\, R_D^{-1/2} \tag{14}$$

olarak tanımlanan T_1 matrisi T ile aynı özdeğerlere sahiptir. Bu durumda $\max(|\lambda_{T1}|) < 1$ olduğunu göstermek, Gauss-Seidel algoritmasının pozitif tanımlı simetrik matrisler için yakınsadığını göstermek için yeterlidir. Bu amaçla

$$R_{L1} = R_D^{-1/2}\, T\, R_D^{-1/2} \tag{15}$$

tanımlamasını yaptığımızda T_1 matrisini

$$T_1 = R_D^{1/2}\, T\, R_D^{-1/2} = -(I + R_{L1})^{-1} R_{L1}^T \tag{16}$$

olarak yazabiliriz. T ile T_1 matrisleri aynı özdeğerlere sahip olduğundan dolayı T_1 matrisinin özdeğerlerini bulmak için

$$-(I + R_{L1})^{-1} R_{L1}^T\, x = \lambda\, x \tag{17}$$

eşitliğini yazabiliriz. Bu eşitliğin her iki tarafını soldan $(I + R_{L1})$ ile çarptığımızda

$$-R_{L1}^T x = \lambda(I + R_{L1})x \tag{18}$$

eşitliğini yazabiliriz. Bu eşitliğin her iki tarafını soldan x^T ile çarparak $x^T x = 1$ özelliğini kullandığımızda

$$-x^T R_{L1}^T x = \lambda\left(1 + x^T R_{L1} x\right) \tag{19}$$

eşitliğini elde ederiz. Buradan λ'yı

$$\lambda = \frac{-x^T R_{L1}^T x}{\left(1 + x^T R_{L1} x\right)} \tag{20}$$

olarak elde ederiz. Burada $x^T R_{L1} x = a$ şeklinde skaler bir sayıdır. Burada

$$a^T = \left(x^T R_{L1} x\right)^T = x^T R_{L1} x = a \tag{21}$$

özelliğini kullanırsak,

$$\lambda^2 = \frac{a^2}{1 + 2a + a^2} \tag{22}$$

sonucu elde edilir. Burada $|\lambda| < 1$ olması için

$$1 + 2a = 1 + x^T R_{L1} x + x^T R_{L1}^T x > 0 \tag{23}$$

olması gerekir. Burada R_{L1} pozitif tanımlı bir matris olduğundan dolayı bu şart her zaman sağlanır. Bu şart her zaman geçerli olduğundan dolayı $|\lambda_k| < 1$, $(k = 1,2,\ldots,M)$ şartı da her zaman geçerlidir.

3.2. Stokastik Yakınsama Analizi

Korelasyon matrisi ve korelasyon vektörü tahminlerinin (7) ve (8) eşitliklerindeki gibi elde edilen verileri kullanarak ardışıl olarak güncellendiğini varsayalım. Burada korelasyon matrisinin başlangıç aşamasında pozitif tanımlılığını garantilemek amacıyla başlangıç değeri $R(0) = \delta.I$ olarak alınabilir. Burada $\delta = 0.1, 0.01, 0.001$, v.s. gibi küçük bir sayı olarak seçilir. Böylece hem başlangıç aşamasında korelasyon matrisinin pozitif tanımlılığı, yani algoritmanın yakınsaması garantilenmiş olur, hem de parametre tahminlerinin başlangıç aşamasındaki yakınsama hızı kontrol edilmiş olur. Bu durumda (9) eşitliğiyle verilen iteratif algoritmayı

$$w(n+1) = -\left(R_L(n) + R_D(n)\right)^{-1} R_U(n) w(n) + \left(R_L(n) + R_D(n)\right)^{-1} p(n) \tag{24}$$

şeklinde yazabiliriz. Burada n hem iterasyon adımını, hem de zaman indisini göstermektedir. Yani iki iteratif işlem birleştirilmiştir. Burada

$$-(R_D + R_L)^{-1} R_U = I - (R_D + R_L)^{-1} R \tag{25}$$

eşitliğinden yararlanarak aynı eşitliği

$$w(n+1) = \left(I - \left(R_D(n) + R_L(n)\right)^{-1} R(n)\right) w(n) + \left(R_D(n) + R_L(n)\right)^{-1} p(n) \tag{26}$$

olarak yazabiliriz. Eşitliğin sağ tarafında bazı düzenlemeler yaptığımızda ise aynı formülü

$$w(n+1) = w(n) + \left(R_D(n) + R_L(n)\right)^{-1} \left[p(n) - R(n)\, w(n)\right] \tag{27}$$

şeklinde yazabiliriz. Parametre hata vektörünü

$$\varepsilon(n) = w(n) - w_0 \tag{28}$$

olarak tanımladığımızda ve yukarıdaki denklemin her iki tarafından w_0'ı çıkararak $w(n)$ değerini parametre hata vektörüne bağlı olarak

$$w(n) = \varepsilon(n) + w_0 \tag{29}$$

şeklinde yazdığımızda aşağıdaki eşitlikleri elde ederiz:

$$(w(n+1) - w_0) = (w(n) - w_0) + \left(R_D(n) + R_L(n)\right)^{-1} \left[p(n) - R(n)(\varepsilon(n) + w_0)\right] \tag{30}$$

$$\varepsilon(n+1) = \varepsilon(n) + \left(R_D(n) + R_L(n)\right)^{-1}\left[p(n) - R(n)\left(\varepsilon(n) + w_0\right)\right]. \tag{31}$$

Eşitliğin sağ tarafını düzenlediğimizde

$$\varepsilon(n+1) = \left(I - \left(R_D(n) + R_L(n)\right)^{-1} R(n)\right)\varepsilon(n) + \left(R_D(n) + R_L(n)\right)^{-1}\left(p(n) - R(n)\, w_0\right) \tag{32}$$

elde edilir. Bu denklem, parametre hata vektörüne bağlı olan ve sistem matrisi

$$I - \left(R_D(n) + R_L(n)\right)^{-1} R(n) \tag{33}$$

olan bir stokastik fark denklemidir. Burada amacımız parametre hata vektörünün beklenen değerinin

$$E\left(\varepsilon(n)\right) \cong 0 \tag{34}$$

şeklinde zamanla sıfıra yakınsadığını göstermektir. Bu aşamada Haykin (1991) tarafından yapılan yaklaşıklığa benzer bir yaklaşıklık yapabiliriz. Haykin (1991) tarafından, stokastik eğim yöntemi olarak da bilinen LMS (**L**east **M**ean **S**quares) algoritmasının stokastik yakınsama analizinde, LMS algoritmasının parametre hata vektörüne ait stokastik fark denkleminin sistem matrisi aşağıdaki gibi verilmiştir:

$$I - \mu \underline{u}(n)\underline{u}(n)^T . \tag{35}$$

Burada μ, LMS algoritmasındaki adım büyüklüğüdür ve $\underline{u}(n)$ vektörü (4) eşitliğindeki gibi tanımlanmıştır. Haykin (1991) tarafından, LMS algoritmasının parametre hata vektörüne ait stokastik fark denkleminin çözümünün, μ adım boyutunun küçük olması durumunda, sistem matrisi

$$E\left(I - \mu \underline{u}(n)\underline{u}(n)^T\right) = I - \mu R \tag{36}$$

olan fark denklemine yakın olduğu gösterilmiştir. Burada Haykin (1991) tarafından

$$E\left(\underline{u}(n)\underline{u}(n)^T\right) \cong R \tag{37}$$

yaklaşıklığı yapılmıştır. Bu durumda Gauss-Seidel algoritmasında rahatlıkla

$$E\left(R(n)\right) \cong R \ , \ E\left(R_D(n)\right) \cong R_D \ , \ E\left(R_L(n)\right) \cong R_L \ \text{ve} \ E\left(p(n)\right) \cong p \tag{38}$$

yaklaşıklıklarını yapabiliriz. Çünkü Gauss-Seidel algoritmasında LMS algoritmasındaki gibi korelasyon matrisinin ve korelasyon vektörünün anlık değeri değil, zaman ortalaması alınmış değeri kullanılmaktadır. Yani, Gauss-Seidel algoritmasının stokastik yakınsama analizinde yaptığımız (38) eşitliğindeki yaklaşıklık, LMS algoritmasının stokastik yakınsama analizinde yapılan (37) eşitliğindeki yaklaşıklığa göre çok daha güçlü bir yaklaşıklıktır. Bu durumda Gauss-Seidel algoritmasının parametre hata vektörünü içeren (32) eşitliğiyle verilen stokastik fark denkleminin her iki tarafının beklenen değerini aldığımızda

$$E\left(\varepsilon(n+1)\right) = E\left(I - \left(R_D(n) + R_L(n)\right)^{-1} R(n)\right)E\left(\varepsilon(n)\right) + E\left(\left(R_D(n) + R_L(n)\right)^{-1}\right)E\left(p(n) - R(n)\, w_0\right) \tag{39}$$

elde edilir. Burada (38) eşitliğini göz önüne alarak

$$\begin{aligned} E\left(p(n) - R(n)\, w_0\right) &\cong E\left(p(n)\right) - E\left(R(n)\right) w_0 \\ &\cong p - R\, w_0 = 0 \end{aligned} \tag{40}$$

yaklaşıklığını yapabiliriz. Ayrıca LMS algoritmasındaki (36) eşitliğine benzer şekilde ve LMS algoritmasındakine göre daha yüksek bir doğruluk ile

$$E\left(I - \left(R_D(n) + R_L(n)\right)^{-1} R(n)\right) \cong I - \left(R_D + R_L\right)^{-1} R \tag{41}$$

yaklaşıklığını rahatlıkla yapabiliriz. Bu durumda (32) eşitliğiyle verilen stokastik fark denkleminin beklenen değerinin yakınsamasını gösteren (39) eşitliğini yüksek bir yaklaşıklık ile

$$E\left(\varepsilon(n+1)\right) = \left(I - \left(R_D + R_L\right)^{-1} R\right)E\left(\varepsilon(n)\right) \tag{42}$$

olarak yazabiliriz. Burada

$$v(n) = E\left(\varepsilon(n)\right) \tag{43}$$

tanımlamasını yaptığımızda aynı fark denklemini

$$v(n+1) = \left(I - \left(R_D + R_L\right)^{-1} R\right)v(n) \tag{44}$$

olarak yazabiliriz. Bu fark denklemini başlangıç değerine bağlı olarak

$$v(n+1) = \left(I - (R_D + R_L)^{-1} R\right)^{n+1} v(0) \tag{45}$$

şeklinde yazabiliriz. Parametre tahminlerinin ortalamasının gerçek değerine yakınsayabilmesi için parametre tahmin hatasının ortalamasının yakınsamasını gösteren bu fark denkleminin sıfıra yakınsaması gerekir. Bunun için, $I - (R_D + R_L)^{-1} R$ matrisinin $(n+1)$. kuvvetinin sıfır olması gerekir. Bu da $I - (R_D + R_L)^{-1} R$ matrisinin bütün özdeğerlerinin 1'den küçük olması şartını gerektirir. Burada R korelasyon matrisi ile bu matrisin köşegen elemanlarını içeren R_D kare matrisi ve alt üçgen kısmını içeren R_L kare matrisi pozitif tanımlı olduğundan, (23) eşitliği gereği bu şart sağlanmaktadır.

4. SONUÇLAR

Bu makalede, Gauss-Seidel algoritmasın adaptif filtre parametrelerini iteratif olarak güncellemek için kullanılması durumunda, hesaplanan parametre tahminlerinin deterministik ve stokastik anlamda yakınsama analizleri yapılmıştır. Parametre tahminlerinin deterministik anlamda yakınsaması için korelasyon matrisinin pozitif tanımlı olmasının yeterli olduğu gösterilmiş, özellikle stokastik anlamda parametre tahmin hatasının beklenen değerinin sıfıra yakınsaması için gerekli şart elde edilerek bu şartın her durumda sağlandığı gösterilmiştir.

5. KAYNAKLAR

1. Albu, F. and Kwan H. K. (2004) Fast block exact Gauss-Seidel pseudo affine projection algorithm, *Electronics Letters*, 40(22), 1451-1452.
2. Bouchard, M. and Albu, F. (2005) The Gauss-Seidel fast affine projection algorithm for multichannel active noise control and sound reproduction systems, *International Journal of Adaptive Control and Signal Processing*, 19(2-3), 107-123.
3. Farhang-Boroujeny, B. (1998) *Adaptive Filters: Theory and Applications*, John Wiley & Sons, Chicester.
4. Golub, G. H. and Van Loan, C. F. (1996) *Matrix Computations*, 3rd Ed., John Hopkins University Press, Baltimore and London.
5. Haykin, S. (1991) *Adaptive Filter Theory*, 2nd Ed., Prentice-Hall, Englewood Cliffs, New Jersey.
6. Koçal, O. H. (1998) A new approach to least squares adaptive filtering, *IEEE International Symposium on Circuits and Systems*, Monterey, California, 261-264.
7. Ng, T. M., Farhang-Boroujeny, B. and Garg, H. K. (2003) An accelerated Gauss-Seidel method for inverse modeling, *Signal Processing*, 83(3), 517-529.