論文 96-21-9-4

4차 고다드 알고리즘의 몇 가지 수렴 성질 : 1. 수렴속도

正會員 최 진 호*, 배 진 수**, 송 익 호**, 박 래 홍***, 박 정 순**

Some Convergence Properties of Godard's Quartic Algorithm : 1. The Rate of Convergence

Jinho Choi*, Jinsoo Bae**, Iickho Song**, Rae-Hong Park***, and Jeongsoon Park** *Regular Members*

요 약

이 논문에서는 눈 먼 균등화에 쓰이는 4차 고다드 알고리즘의 수렴에 대하여 살펴보았다. 평균 성능 함수가 가장 작은 값을 갖는 점 부근에서의 4차 고다드 알고리즘의 국소적 성질을 이해함으로써 적응이득을 결정하였다. 4차 고다드 알고리즘의 정규화된 적응이득과 수렴속도는 결정 지향 알고리즘과 견주어 볼 때 적응이득은 작아야 한다는 것과 수렴속도는 빠르다는 것을 보였다.

ABSTRACT

Convergence analysis on Godard's quartic (GQ) algorithm used for blind equalization is accomplished in this paper. The first main result is an explanation of the local behavior of the GQ algorithm around the global minimum point of the average performance function, from which we can determine the adaptation gain. It is shown that the normalized adaptation gain of the GQ algorithm should be smaller than that of the decision directed (DD) algorithm. In addition, it is observed that the GQ algorithm converges faster than the DD equalization algorithm.

I. 머리말

학습 신호가 없을 때에도 쓸 수 있는 여러 가지 적응 여파 알고리즘들이 개발되어 왔는데 이들은 채널 균등화나 신호 복원과 같은 곳에 쓸 수 있다. 눈 먼 균등화기도 (blind equalizer) 이런 적응 여파기의 한 갈래이다. 눈 먼 균등화 알고리즘의 성능 함수는 일반적으로 극소점을 몇 개 갖는 꼴이다. 어떤 성능 함수들에 대해 안정된 극소점들이 존재함이 증명되어 있으나, 일반적으로 눈 먼 균등화 알고리즘이 언제나 수렴하는 것은 아니다 [1-3]. 한편, 알려진 입력 신호를 쓰는 균등화기의 성능 함수는 볼록 함수이며 그

*데이콤 종합연구소 PCS 연구팀
** 한국과학기술원 전기 및 전자공학과
*** 서강대학교 전자공학과
論文番號 : 95264-0807
接受日字 : 1995年 8月 7日

성능 함수는 언제나 최소값으로 수렴한다. 요컨대, 눈 먼 균등화 알고리즘의 단점은 뒤떨어지는 수렴 성질이다.

적응 여파기의 출력을 $z(k)$라 하고 성능 함수

$$J(k)=\frac{1}{4}\ (z^2(k)-\gamma)^2 \tag{1}$$

을 생각해보자. 여기서, γ는 입력 신호의 통계량으로 결정한다. 또한 기울기 경사법을 쓰는 여파기에서 매개변수는 다음 식을 써서 결정할 수 있다.

$$\theta(k+1)=\theta(k)-\mu\frac{\partial J(k)}{\partial\theta(k)}. \tag{2}$$

여기서, μ는 적응 이득이다. 이것이 눈 먼 균등화기를 위한 4차 고다드 알고리즘이다. 이 알고리즘은 눈 먼 균등화에 많이 쓰인다. 식 (2)에 나타난 알고리즘의 수렴에 관한 성질들은 부분적으로 [2-4]에서 다루어졌다.

적응 알고리즘의 수렴 성질은 적응 이득 μ의 영향을 받는다. 만약 μ가 크면 적응 속도는 빨라질 것이다. 그러나, μ가 너무 크면 적응 알고리즘이 발산하므로 이 값을 크게 하는 데에는 한계가 있다. 덧붙여 말하면, 수렴 속도는 입력 신호 공분산 행렬의 고유값 분포의 영향도 받는다.

이 논문에서는 4차 고다드 알고리즘이 PAM신호의 눈 먼 균등화에 쓰일 때에 4차 고다드 알고리즘의 두 가지 주된 성질을 살펴보았다. 곧, 평균 성능 함수가 가장 작아지는 점 부근에서의 국소 성질과 흡인 영역에서의 적응 이득의 상한값, 수렴속도, 안정성을 살펴보았다.

Ⅱ. 평균 성능 함수의 성질

잡음이 없을 때, 충격응답이 $\{h_i,\ i=0,\ 1,\dots\}$인 선형 채널의 입력 $\{a(i),\ i=1,\ 2,\dots\}$이 독립이고 같은 확률을 가지는 PAM 신호라면 그 출력은

$$y(k)=\sum_{i=0}^{\infty}h_i a(k-i) \tag{3}$$

이다. 균등화기의 매개변수 벡터가 $\theta(k)=[\theta_0(k),\ \theta_1(k),\dots]^T$ 라면 균등화기의 출력은

$$z(k)=\sum_{i=0}^{\infty}\theta_i(k)\,y(k-i) \tag{4}$$

이다. 적응 알고리즘 (2)를 쓸 때 γ를

$$\gamma=\frac{E\{a^4(k)\}}{E\{a^2(k)\}} \tag{5}$$

로 잡고 매개변수 벡터 $\theta(k)$의 처음 값을 최적 매개변수 θ^*과 충분히 가깝게 잡으면 부호 사이 간섭을 없애주는 벡터 θ^*을 얻을 수 있다. 곧, $\theta(k)=\theta^*$이면 성능 함수 $J(k)$는 가장 작은 값을 갖게되고 시간 지연이 있을 수는 있지만 균등화기의 출력은 입력신호와 비슷하게 된다 [1]. 탭이 무한히 많고 비용함수가 완벽할 때에만 $a(k)=z(k)$를 얻을 수 있으며, 일반적으로는 오차가 있기 마련이다. 여기서는 시간 지연은 없다고 가정한다.

평균 성능 함수를 $\overline{J_0}=E\{J(k)\}$라 하고, (2)의 평균을 얻으면,

$$\theta(k+1)=\theta(k)-\mu\frac{\partial\overline{J_0}}{\partial\theta}\Big|_{\theta=\theta(k)} \tag{6}$$

임을 알 수 있다. 여기서,

$$\frac{\partial\overline{J_0}}{\partial\theta}\Big|_{\theta=\theta(k)}=E\{(z^2(k)-\gamma)\,z(k)\,Y(k)\} \tag{7}$$

이고 $Y(k)=[y(k),\ y(k-1),\dots]^T$이다. 한편, $z(k)=Y^T(k)\theta(k)$이므로 (6)과 (7)에서

$$\begin{aligned}\theta(k+1)&=\theta(k)-\mu E\{(z^2(k)-\gamma)Y(k)Y^T(k)\}\theta(k)\\&=\{I-\mu E[(z^2(k)-\gamma)Y(k)Y^T(k)]\}\theta(k)\end{aligned} \tag{8}$$

을 얻을 수 있다. 이제, $\theta(k)=\theta^*$일 때는 $z(k)\approx a(k)$이고 $\partial\overline{J_0}/\partial\theta|_{\theta(k)=\theta^*}=0$이므로

$$E[a^2(k)Y(k)Y^T(k)]\theta^*=\gamma E[Y(k)Y^T(k)]\theta^* \tag{9}$$

이다. 행렬 $E[Y(k)Y^T(k)]$와 $E[a^2(k)Y(k)Y^T(k)]$를 각각 R_y, B_y라고 하면

$$E[a^2(k)a(k-m-p)a(k-n-q)]=\begin{cases}E[a^4(k)], & m+p=n+q=0\text{일 때},\\ E^2[a^2(k)], & m+p=n+q\neq 0\text{일 때};\\ 0, & \text{그 밖}.\end{cases} \tag{10}$$

이므로 B_y의 $(m+1, n+1)$째 원소는 다음과 같다.

$$b(m, n)=E[a^2(k)y(k-m)y(k-n)]$$

$$=\sum_{p=0}^{\infty}\sum_{q=0}^{\infty}h_p h_q E[a^2(k)a(k-m-p)a(k-n-q)] \quad (11)$$

$$=\begin{cases} h_0^2 E[a^4(k)]+\sum_{p=1}^{\infty}h_p^2 E^2[a^2(k)], \ m=n=0\text{일때}; \\ \sum_{p=0}^{\infty}h_p h_{p+m-n}E^2[a^2(k)], \ m\neq 0 \text{ or } n\neq 0\text{일때}. \end{cases}$$

마찬가지로, R_y의 $(m+1, n+1)$째 원소 $r(m, n)$은

$$r(m, n)=\sum_{p=0}^{\infty}h_p h_{p+m-n}E[a^2(k)]. \quad (12)$$

여기서,

$$\delta \equiv E^2[a^2(k)]-E[a^4(k)] \quad (13)$$

라 두면, 다음 식을 얻을 수 있다.

$$b(m, n)-\gamma r(m, n)=\begin{cases} \delta(\sum_{p=0}^{\infty}h_p^2-h_0^2), \ m=n=0\text{일때}; \\ \delta\sum_{p=0}^{\infty}h_p h_{p+m-n}, \text{ 그밖}. \end{cases}$$

식 (10)과 (12)를 쓰면

$$B_y-\gamma R_y=\delta(R_h-h_0^2 I_0). \quad (14)$$

여기서, $[R_h]_{m+1, n+1}=\sum_{p=0}^{\infty}h_p h_{p+m-n}$이고

$I_0=\begin{bmatrix} 1 & 0 & 0 & \cdots \\ 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}$ 이다. 이때 입력 신호 $a(k)$가 모든 k에 대해 일정한 크기를 갖는다면, (14)의 δ는 0이 되어

$$B_y=\gamma R_y. \quad (15a)$$

곧

$$b(m, n)=\gamma r(m, n). \quad (15b)$$

그렇지 않으면 δ는 일반적으로 0이 아니므로 (9)를 써서 (14)를 바꾸면

$$R_h\theta^*=h_0^2\begin{bmatrix} \theta^* \\ 0 \\ 0 \\ \vdots \end{bmatrix} \quad (16)$$

여기서 θ_0^*는 θ^*의 첫째 원소이다. 이는 평균 성능 함수 $\overline{J_0}$가 가장 작은 값을 갖도록 하는 θ^*에 대한 필요조건이다. 또한 (16)에서 최적 매개변수 벡터 θ^*를 갖는 균등화기는 백색화 작용을 한다는 것을 알 수 있다.

Ⅲ. 4차 고다드 알고리즘의 안정성

식 (6)의 매개변수 벡터 $\theta(k)$가 최적 매개변수 벡터 θ^*와 비슷한 값을 갖는다면 [5]

$$\frac{\partial}{\partial\theta}\overline{J_0}|_{\theta=\theta(k)}=$$

$$\frac{\partial}{\partial\theta}\overline{J_0}|_{\theta=\theta^*}+M(\theta^*)(\theta(k)-\theta^*)+O(\|\theta(k)-\theta^*\|^2). (17)$$

여기서, $M(\theta^*)=\frac{\partial^2}{\partial\theta^2}\overline{J_0}|_{\theta=\theta^*}$이고 $\|\cdot\|$은 벡터의 크기를 나타낸다. 이어서, 식 (6)을 쓰면

$$\theta(k+1)=\theta(k)-\mu M(\theta^*)(\theta(k)-\theta^*)+O(\|\theta(k)-\theta^*\|^2). \quad (18)$$

여기서, $\tilde{\theta}(k)\equiv\theta(k)-\theta^*$라고 하고, $O(\|\theta(k)-\theta^*\|^2)$항을 무시하면

$$\tilde{\theta}(k+1)\simeq\tilde{\theta}(k)-\mu M(\theta^*)\tilde{\theta}(k). \quad (19)$$

여기에

$$M(\theta^*)=3B_y-\gamma R_y \quad (20)$$

을 대입하면 [3],

$$\tilde{\theta}(k+1)\simeq\tilde{\theta}(k)-\mu(3B_y-\gamma R_y)\tilde{\theta}(k). \quad (21)$$

식 (21)을 써서 (18)의 안정성을 점검할 수 있다. 곧,

(21)의 차분 방정식의 풀이 $\tilde{\theta}(0)$가 0으로 수렴한다면 (18)의 차분방정식을 만족시키는 풀이도 $\tilde{\theta}(0)$이 충분히 작으면 0으로 수렴하게 된다 [5].

따라서, 이제부터는 (21)의 안정성만을 생각하면 된다. $\delta=0$이면

$$3B_y-\gamma R_y=2B_y=2E[a^2(k)Y(k)Y^T(k)]$$

이 되어 헤시안 $M(\theta^*)$는 반양정치이다 (positive semi-definite). 그리고, $\delta \neq 0$일 때에도 (10)과 (12)에서

$$3B_y-\gamma R_y=\delta_3 R_h+\tau I_0. \tag{22}$$

여기서, $\delta_3 \equiv 3E^2[a^2(k)]-E[a^4(k)]$이고 $\tau \equiv -3\delta h_0^2$이다. 이 때에도 δ_3이 음이 아니면 $\delta \leq 0$이고 행렬 R_h와 I_0가 모두 반양정치이므로 헤시안 $M(\theta^*)$은 반양정치이다. 곧,

$$\begin{aligned} M(\theta^*) &= \delta_3 R_h+\tau I_0 \\ &\geq 0. \end{aligned} \tag{23}$$

다시 말해서, (23)을 만족시키거나 $\delta_3 \geq 0$이면 알고리즘 (21)이 최적점 부근에서 국소적 안정성을 갖게 된다 [3, 4]. 그런데, $\delta_3 \geq 0$일 조건은

$$2E^2[a^2(k)] \geq \mathrm{var}[a^2(k)]$$

이므로 $a(k)$의 분산이 작아야 한다. 바꾸어 말하면, 4차 고다드 알고리즘이 안정적이려면 신호 크기가 가질 수 있는 값들이 너무 퍼져 있으면 안된다. 보기를 들면, $a^2(k) \in \{1, 16, 256, 4096\}$이고 동일한 분포를 갖는다면, δ_3는 음수가 된다. 그런데, 실제상황에서는 인접한 두 신호사이의 거리는 일반적으로 같고 δ_3는 양의 값을 갖게 된다. 따라서, 대부분의 실제 신호에 대해 (18)의 안정성을 보장할 수 있다.

Ⅳ. 4차 고다드 알고리즘의 수렴속도

적응 알고리즘에서는 적응 이득 μ가 수렴속도와 안정도를 결정하는데, 성능 함수가 볼록 함수라 할지라도 μ가 너무 크면 적응 알고리즘이 발산하므로 적응 이득 μ의 상한값은 적응 알고리즘이 수렴하도록 정해져야 한다.

식 (21)과 (22)에서 적응 알고리즘이 수렴하도록 하는 μ의 범위는 다음과 같다.

$$0 < \mu < \frac{2}{\delta_3 \lambda_{\max}^G}. \tag{24}$$

여기서, $\lambda_{\max}^G$는 $\hat{\tau}=\tau/\delta_3$일 때 행렬 $R_h+\hat{\tau}I_0$의 고유값 가운데에서 가장 큰 값이다.

4차 고다드 알고리즘의 수렴속도를 결정 지향 (decision directed) 알고리즘과 견주어 보자 [6]. 결정지향 알고리즘의 성능 함수는

$$J_D(k)=\frac{1}{2}\{z(k)-dec(z(k))\}^2. \tag{25}$$

여기서, $dec(\cdot)$는 결정 장치이다. 이 알고리즘이 수렴하려면 다음 조건이 만족되어야 한다.

$$0 < \mu < \frac{2}{E[a^2(k)]\,\lambda_{\max}^D}. \tag{26}$$

여기서, $\lambda_{\max}^D$는 행렬 R_h의 가장 큰 고유값이다.

헤시안 $M(\theta^*)$의 가장 큰 고유값 $\lambda_{\max}^M$은 [7]

$$\begin{aligned} \lambda_{\max}^M &= \delta_3 \lambda_{\max}^G \\ &\geq \delta_3 \lambda_{\max}^D \end{aligned} \tag{27}$$

을 만족시킨다. 결정 지향 알고리즘과 4차 고다드 알고리즘을 각각 $E[a^2(k)]$과 δ_3을 써서 정규화하면, 4차 고다드 알고리즘의 정규화된 적응 이득은 결정 지향 알고리즘의 정규화된 적응 이득보다 항상 작게 된다. 이것은 특별한 경우에 대해서 [2]에서도 다루어졌다.

· 보기 1

값 $\{-3, -1, 1, 3\}$을 같은 확률로 갖는 PAM 시스템을 생각해 보자. AR(1)채널을

$$y(k)+0.6y(k-1)=a(k)$$

라 하면 $\delta_3=34$, $\lambda_{\max}^G=3.4$, $\lambda_{\max}^D=2.5$이다. 곧, 4차 고다드 알고리즘의 적응 이득이 결정 지향 알고리즘의 적응 이득 보다 작아야 알고리즘이 수렴한다.

행렬 $M(\theta^*)$과 R_y의 고유값 분포를 각각 S_G와 S_D라고 하면 다음 두 식을 얻을 수 있다.

$$S_G = \frac{\lambda_{\max}^D + \alpha\hat{\tau}}{\lambda_{\min}^D + \beta\hat{\tau}} \tag{28}$$

$$S_D = \frac{\lambda_{\max}^D}{\lambda_{\min}^D}. \tag{29}$$

여기서, $\lambda_{\max}^D$은 행렬 R_h의 가장 작은 고유값이며 $\alpha, \beta \geq 0$, $\alpha + \beta \leq 1$이다 [7]. 또

$$\eta \equiv \frac{S_D}{S_G} \tag{30}$$

라 두면, $\eta > 1$일 때에는 4차 고다드 알고리즘의 고유값 분포가 결정 지향 알고리즘의 고유값 분포 보다 더 작게 되고 따라서 수렴 속도가 더 빠르게 된다. 그러므로, 이제 $\eta < 1$을 증명하여 4차 고다드 알고리즘이 더 빠른 수렴속도를 가지고 있음을 증명하고자 한다. 식 (28)-(30)에서

$$\eta = \frac{\lambda_{\max}^D}{\lambda_{\min}^D} \cdot \frac{\lambda_{\min}^D + \beta\hat{\tau}}{\lambda_{\max}^D + \alpha\hat{\tau}} \tag{31}$$

이고 $\alpha, \beta \geq 0$과 $\alpha + \beta \leq 1$을 이용하면

$$\frac{1}{1 + (\hat{\tau}/\lambda_{\max}^D)} \leq \eta \leq 1 + \frac{\hat{\tau}}{\lambda_{\min}^D} \tag{32}$$

이다. 그리고 $\lambda_{\max}^D \gg \hat{\tau}$이라면

$$\eta \approx 1 + \beta\frac{\hat{\tau}}{\lambda_{\min}^D} \tag{33}$$

을 얻을 수 있다. 곧, $\lambda_{\max}^D \gg \hat{\tau}$일 때에는 4차 고다드 알고리즘의 수렴속도는 결정지향 알고리즘의 수렴 속도보다 빠르다는 것을 알 수 있다. 또한, S_D의 값이 매우 작고 $\lambda_{\min}^D \gg \hat{\tau}$이면 두 알고리즘의 수렴 속도는 같다.

· 보기 2

보기 1과 같은 채널에서 $S_D = 4.0$이다. 한편, S_G의 값은 입력 신호의 통계적 특성으로 결정되는데, 보기 1의 4 PAM신호에 대한 S_G의 값은 3.14이고, $a(k)$가 1과 −1 두 값만을 가지는 PAM 신호라면 S_G의 값은 4.0이다.

V. 맺음말

이 논문에서는 고다드 알고리즘을 쓰는 눈 먼 등화기의 수렴 성질이 채널의 특성과 상관이 있다는 것을 밝혔다. 또 알고리즘이 발산하지 않는 적응 이득의 범위를 구할 수 있음과 그 정규화된 수렴속도와 적응 이득은 결정지향 알고리즘과 견주어 볼 때 작다는 것을 보였다. 이 논문과 이어지는 2부에서는 4차 고다드 알고리즘의 평균 성능 함수의 기하에 대해서 살펴보고자 한다.

참 고 문 헌

1. D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. Comm.*, vol. COM-28, pp. 1867-1875, November 1980.
2. G. J. Foschini, "Equalizing without altering or detecting data," *AT&T Tech. Jour.*, vol. 64, pp. 1885-1911, October 1985.
3. Z. Ding, R. A. Kennedy, B. D. O. Anderson, and C. R. Johnson, "Ill-convergence of Godard blind equalizers in data communication systems," *IEEE Trans. Comm.*, vol. COM-39, pp. 1313-1327, September 1991.
4. C. R. Johnson, Jr., S. Dasgupta, and W. A. Sethares, "Averaging analysis of local stability of a real constant modulus algorithm adaptive filter," *IEEE Trans. Acoust., Speech, Signal Proc.*, vol. ASSP-36, pp. 900-910, June 1988.
5. J. E. Slotine and W. Li, *Applied Nonlinear Control*, Prentice-Hall, Englewood Cliffs, NJ, 1991.
6. O. Macchi and E. Eweda, "Convergence analysis of self-adaptive equalizers," *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 161-176, March 1984.
7. G. H. Golub and C. F. Van Loan, *Matrix Computation*, The Johns Hopkins University Press, Baltimore, MD, 1983.

최 진 호(Jinho Choi) 정회원

1967년 1월 14일생

1989년 2월 : 공학사, 서강대학교 전자공학과

1991년 2월 : 공학석사, 한국과학기술원 전기및전자공학과

1994년 2월 : 공학박사, 한국과학기술원 전기및전자공학과

1994년 2월~현재 : 데이콤 종합연구소 선임연구원

배 진 수(Jinsoo Bae) 정회원

한국통신학회 논문지 제20권 제12호 참조

현재 : 한국과학기술원 전기및전자공학과 박사과정

송 익 호(Iickho Song) 정회원

한국통신학회 논문지 제21권 제1호 참조

현재 : 한국과학기술원 전기및전자공학과 부교수

박 래 홍(Rae-Hong Park) 정회원

한국통신학회 논문지 제19권 제5호 참조

현재 : 서강대학교 전자공학과 교수

박 정 순(Jeongsoon Park) 정회원

1970년 11월 17일생

1996년 2월 : 연세대학교 전파공학과 공학사

1996년 3월~현재 : 한국과학기술원 전기 및 전자공학과 석사과정

※주관심 분야 : 통계학적 신호처리, 신호검파