

순위 통계량으로 확률 신호를 검파하는 방법 : 제1부. 한 표본을 쓸 때

正會員 宋 翊 鎬* 正會員 吳 澤 相* 正會員 嚴 泰 相* 正會員 韓 永 玉*

Methods of Random Signal Detection with Rank Statistics : Part I. The One-Sample Case

lick Ho SONG*, Taek Sang OH*, Tae Sang UHM*, Young Ok HAN* *Regular Members*

要 約 가산성 잡음이 섞여 있는 확률 신호를 순위 통계량에 바탕을 두고 검파하는 방식을 얻었다. 이 검파기가 확률 신호 국소 최적 검파기와 비슷하다는 것과 알려진 신호 국소 최적 순위 검파기를 일반화한 것이라는 것을 보였다. 뿐만 아니라 이 검파기의 성능을 다른 검파기들과 견주어 보았다.

ABSTRACT A detection scheme which has bases in rank statistic is obtained for detection of random signals in additive noise. It is shown that the detector has similarities to the locally optimum detector for random signals and that it is a generalization of the locally optimum rank detector for known signals. Performance of the detector is also considered together with that of other detectors.

I. 머리말

잡음이 섞인 약한 신호를 검파하는 문제는 여러 방향으로 연구되어 왔다: 가산성 잡음이 있을 때 알려진 신호(known signal)와 확률 신호(random signal)를 검파하는 국소 최적(locally optimum, LO) 검파기 연구¹⁾, 알려진 신호를 검파하는 국소 최적 순위(locally optimum rank, LOR) 검파기 연구²⁾, 일반화된 관측 모형에서(generalized observation model) 국소 최적 검파기 연구³⁾, 복합 신호(composite signal) 모형에서 국소 최적 검파기 연구가⁴⁾ 이의 대표적인 보기이다.

잡음 세기에 견주어 큰 세기를 가지고 있는 신호를 검파하는 것보다 신호 세기가 약한 신호(weak signal)를 검파하는 것이 어렵기 때문에 약한 신호 검파에 쓸모 있는 국소 최적 검파기는

신호 검파 문제에서 많은 관심을 끌어 왔다. 또한 국소 최적 검파기는 최적(optimum) 또는 균일 최강(uniformly most powerful, UMP) 검파기에 견주어 상대적으로 구현하기 쉽다는 좋은 점을 가지고 있다. 일반적으로 국소 최적(순위) 검파기의 검정통계량(test statistic)은 알려진 신호 값이나 신호 공분산 함수로 만들어 질 뿐 신호의 다른 특별한 정보를 필요로 하지 않는다는 것이 알려져 있다.

이 논문에서는 통계학에서 널리 쓰이는 순위 검정(rank test) 방법을 응용하여 가산성 잡음이 섞여 있을 때 평균이 0인 약한 확률 신호를 검파하는 문제를 다루겠다. 관측값을 쓰지 않고 관측의 순위(rank)와 부호(sign)를 쓰면 기억 용량을 크게 필요로 하지 않으면서 관측값에 대한 가정이 그다지 많이 바뀌지 않을 때에는 성능이 안정된 검파기를 얻으리라 예상할 수 있다.

*韓國科學技術院 電氣및 電子工學科
Dept. of Electrical Eng., KAIST
論文番號 : 91-26 (接受1991. 1. 10)

II. 전 제

2.1 관측 모형

신호 검파에서 이상 서간 관측 X 의 가산성 잡음 모형(additive noise model)은 다음과 같이 나타낼 수 있다.

$$X_i = \theta V_i + W_i, \quad i=1,2,\dots,n. \quad (1)$$

(1)에서 n 은 입력 경로 하나의 표본 크기(sample size)이고, 곧 관측량의 갯수이고, V_i 는 i 번째 표본의 공통 신호 성분이며, θ 는 신호 세기를 나타내는 매개 변수인데 이것은 0이거나 (귀무가설) 0보다 큰 값을 (대립가설) 갖는다. W_i 는 i 번째 관측량의 가산성 잡음성분이다. 일반적으로 $V_i = c_i + S_i$ 로 나타낼 수 있는데, 여기서 $E\{V_i\} = c_i$ 는 알려진 신호 성분이고, S_i 는 평균이 0이고 공분산 함수가 $r(i, k) = E\{S_i S_k\}$ 인 확률 신호 성분이다. 이 논문에서는 $c_i = 0$ 인 때만을 생각하기로 한다. 잡음 성분 W_i 는 평균이 0, 분산이 σ_i^2 이고 확률 밀도 함수가 f_W 인, 독립이고, 같은 분포를 갖는(independent and identically distributed, i.i.d.), 확률 변수이다. 또한 신호와 잡음은 서로 독립이다.

확률 신호를 검파하는 국소 최적 순위 검파기의 검정 통계량을 얻으려면 몇가지 가정이 필요하다: 잡음 확률 밀도 함수 f_W 는 우함수이고 ($f_W(x) = f_W(-x)$), Fisher 정보가⁽⁵⁾ 유한하며 ($\int_{-\infty}^{\infty} (\frac{f'_W(x)}{f_W(x)})^2 f(x) dx < \infty$), 미분과 적분 같은 여러 수학적 연산의 순서를 바꿀 수 있도록 정규 조건(regularity conditions)을⁽⁶⁾ 만족시키며 충분히 완만하다고 가정하자.

II.2 지난 연구 결과

$V_i = c_i$ 일 때, 곧 $S_i = 0$ 일 때, 알려진 신호를 검파하는 국소 최적 검파기의 검정 통계량 T_{io}^k 와 국소 최적 순위 검파기의 검정 통계량 T_{ior}^{k*} 는 c_i 와 f_W 의 함수로 나타낼 수 있는데 그것들은 다음과 같다⁽¹⁾⁽²⁾.

$$T_{io}^k(X) = \sum_{i=1}^n e_i g(X_i) \quad (2)$$

그리고

$$T_{ior}^{k*}(X) = \sum_{i=1}^n c_i Z_i c_i(Q_i). \quad (3)$$

(2)와 (3)에서 $\mathbf{Z} = (Z_1, Z_2, \dots, Z_n)$ 는 관측 벡터 $\mathbf{X} = (X_1, X_2, \dots, X_n)$ 의 부호 통계량(sign statistic)이고 $\mathbf{Q} = (Q_1, Q_2, \dots, Q_n)$ 는 X 의 크기 순위(magnitude rank) 통계량이며 Q_i (곧, Q_i 는 집합 $M = \{X_1, X_2, \dots, X_n\}$ 에서 X_i 의 순위임).

$$c_i(i) = E\{g(|X[i]|)\} \quad (4)$$

이고

$$g(x) = -\frac{f'_W(x)}{f_W(x)} \quad (5)$$

또한 (4)에서 $|X[i]|$ 는 집합 M 의 i 번째 순서 통계량(order statistic)이다. T_{io}^k 가 국소 최적 비선형성 g 를 지난 관측값을 쓰는 한편 T_{ior}^{k*} 는 점수 함수(score function) c_i 을 지난 관측의 크기 순위를 쓰는 것이 T_{io}^k 와 T_{ior}^{k*} 가 다른 점이다. 또한 T_{ior}^{k*} 에서는 관측 부호통계량도 함께 쓰는 데, 이는 관측 크기 순위를 관측값 대신 쓰는데 따르는 정보의 손실을 메꾸려는 것이라고 볼 수 있다.

한편 $V_i = S_i$ 일 때, 곧 $c_i = 0$ 일 때, 확률 신호를 검파하는 국소 최적 검파기의 검정통계량은⁽³⁾ 다음과 같다.

$$T_{io}(X) = \sum_{i=1}^n \sum_{k=1}^n r(i, k) g(X_i) g(X_k) + \sum_{i=1}^n \sigma_i^2 [h(X_i) - g^2(X_i)] \quad (6)$$

(6)에서

$$h(x) = \frac{f''_W(x)}{f_W(x)} \quad (7)$$

이고 $\sigma_i^2 = r(i, i)$ 는 S_i 의 분산이다.

III. 검정 통계량

이제 이 논문의 주제인, 모형 (1)에서 $T=S$ 일 때 (곧, $\epsilon=0$ 일 때), 화물 신호를 순위 검파하는 문제를 생각하자. 가신성 삽입이 있을 때 화물 신호를 검파하는 문제는 관측의 결합 화물 밀도 함수에 대한 귀무 가설(null hypothesis)과 대립 가설(alternative hypothesis)의 검정 문제로 나타낼 수 있다: 이 때, 귀무가설은

$$H: f(x|H) = \sum f_n(x) \quad (8)$$

이고, 대립가설은

$$K: f(x|K) = \int_{K^n} U f(s) ds \quad (9)$$

이다. (9)에서

$$U = \prod_{i=1}^n f_n(x_i - \theta s_i), \quad (10)$$

$f(s)$ 는 신호 S , $i=1, 2, \dots, n$ 의 결합 화물 밀도 함수이고 R^n 은 n 차원 실수 공간이다. 위 귀무가설과 대립가설을 크기 순위와 부호로 나타내면

$$H: P(\mathbf{q}, \mathbf{z}|H) = \frac{1}{2^n n!} \quad (11)$$

및

$$K: P(\mathbf{q}, \mathbf{z}|K) = \int_B \int_{K^n} U f(\mathbf{s}) d\mathbf{s} d\mathbf{x}, \quad (12)$$

이다. (12)에서 $B = \{\mathbf{x} | \mathbf{Q} = \mathbf{q}, \mathbf{Z} = \mathbf{z}\}$ 이다. 일반화된 Neyman-Pearson 정리를^{6,7)} 쓰면 화물 신호 국소 최적 순위 검파기의 검정 통계량은 다음과 같음을 보일 수 있다.

$$T_{loc}(X) = \sum_{i=1}^n \sum_{k=1}^n r_i(i, k) Z_i Z_k c_2(Q_i, Q_k) + \sum_{i=1}^n \sigma_i^2 [d(Q_i) - c_2(Q_i, Q_i)]. \quad (13)$$

(13)을 얻는 과정을 부록에 나타나 있으며, (13)에서

$$c_2(i, k) = E\{g(X_i | i) g(X_k | k)\} / H \quad (14)$$

이고

$$d(i) = E\{h_i(X_i | i)\} / H, \quad (15)$$

T_{loc} 에서 $g(X_i)g(X_k)$ 를 $c_2(Q_i, Q_k)$ 로, h 를 d 로 바꾸고 부호 통계량을 넣으면 T_{loc} 를 얻을 수 있으므로 두 검정 통계량은 같은 꼴을 갖는다고 할 수 있다. T_{loc} 를 다음과 같이 해석할 수 있다: 관측에 신호 성분과 삽입 성분이 함께 있다고 가정하자. 그러면 관측의 절대값이 커지기 때문에 관측의 크기 순위를 신호 부호에 관계없이 신호 성분이 없을 때 보다 커진다. 다음으로, $r_i(i, k)$ 가 양수이면 Z_i 와 Z_k 가 같은 부호를 갖게 되고 음수이면 서로 다른 부호를 갖는다. 어느 때에나 대립 가설 아래에서는 $r_i(i, k)Z_i Z_k$ 가 커진다. 한편 점수 함수 $c_2(i, k)$ 는 i 또는 k 가 커짐에 따라 커지는 함수이므로⁸⁾ 관측이 신호 성분을 가지고 있으면 관측에 점음만 있을 때보다 검정 통계량이 커진다.

IV. 성능 특성

국소 최적 순위 검파기의 점근 성능과 유한 표본 성능을 잘 알리진 다른 검파기들과 견주어 보자.

4.1 점근 성능

검파기의 점근 성능의 측도로 널리 쓰이는 것 가운데 하나가 점근 상대 효율(asymptotic relative efficiency, ARE)이다. 두개의 검파기 D_1 과 D_2 의 점근 상대 효율 $ARE_{1,2}$ 는 정규 조건이 만족될 때⁹⁾ 다음과 같이 나타낼 수 있다.

$$ARE_{1,2} = \frac{\epsilon_1}{\epsilon_2} \quad (16)$$

(16)에서

$$\epsilon_1 = \lim_{n \rightarrow \infty} \frac{(E\{T|K\})^2}{nV\{T|H\}} \quad (17)$$

는 검파기 D 의 효능(efficacy)이다. (17)에서 θ 은 θ 에 대한 2계 미분을 나타내고, T , $E\{T|K\}$ 와 $V\{T|H\}$ 는 각각 검파기 D 의 검정 통계량, 신호 성분이 있을 때 T 의 평균과 신호 성분이 없을 때 T 의 분산을 나타낸다. 조금 성가신 계산을 몇 단계 거치면 국소 최적 순위 검파기의 효능은 다음과 같음을 보일 수 있다.

$$\epsilon_{1,0k} = 2[\langle r^2 \rangle - \langle \sigma^4 \rangle] I^2(f_u) + [\langle \sigma^4 \rangle - \langle \sigma^2 \rangle^2] I_2(f_u). \quad (18)$$

(18)에서

$$\langle r^2 \rangle \triangleq \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \sum_{k=1}^n r^2(i, k), \quad (19)$$

$$\langle \sigma_j \rangle \triangleq \lim_{n \rightarrow \infty} \frac{1}{n} \sigma_j^2, \quad j=2,4, \quad (20)$$

$$I(f) = \int_{-\infty}^{\infty} \left(\frac{f'(x)}{f(x)} \right)^2 f(x) dx \quad (21)$$

이고

$$I_2(f) = \int_{-\infty}^{\infty} \left(\frac{f''(x)}{f(x)} \right)^2 f(x) dx \quad (22)$$

이다. (18)과 확률 신호 국소 최적 검파기의 효능을⁽¹⁾ 견주어 보면 다음을 쉽게 보일 수 있다.

$$ARE_{LOR10} = 1 - \frac{\langle \sigma^2 \rangle^2 I_2(f_u)}{2[\langle r^2 \rangle - \langle \sigma^4 \rangle] I^2(f_u) + \langle \sigma^4 \rangle I_2(f_u)} \quad (23)$$

4.2. 유한 표본 성능

이제는 국소 최적 순위 검파기와 다른 검파기의 유한 표본 성능을 살펴보자. 일반적으로 검파기의 유한 표본 성능 특성은 주어진 매개 변수값에 따라 바뀌므로 매개변수 값이 달라지면 이미 얻은 결과가 쓸모 없을 수도 있다. 그렇지만 여러 검파기의 성능을 견주어 보는데 유한 표본 성능 비교는 꽤 타당성을 지니고 있다. 한편 실제 응용에서는 표본 크기가 유한하기 때문에 검파기의 유한 표본 성능이 점근 성능보다 실질적인 면에서 더욱 뜻 있다는 것은 쉽게 짐작이 간다.

$n=50$ 일 때 여러 검파기의 검파 확률을 Monte Carlo 방법으로 얻었다. 가산성 잡음 W 는 평균이 0, 분산이 1인 정규 확률 변수라 가정하고 IMSL의 GGNML서브루틴에서 만들었다. 또한 확률 신호 S 는 평균이 0, 분산이 $\frac{2i}{n+1}$ 인 정규 확률 변수(Gaussian random variable)라 가정하여 $\frac{1}{n} \sum_{i=1}^n \sigma_i^2 = 1$ 을 만족시키도록 하였으며, IMSL의 GGNML 서브루틴에서 얻었다.

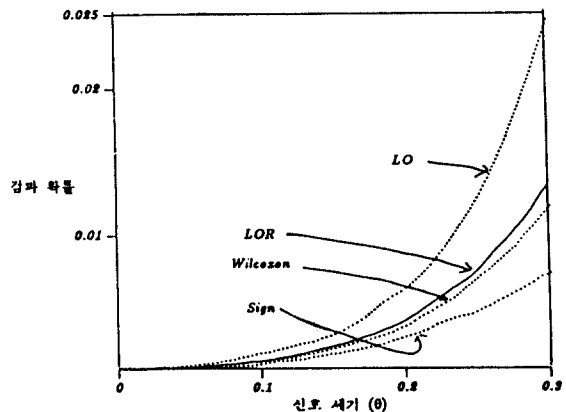


그림 1. 검파기의 성능($P_u=10^{-3}$ 일 때)
Fig. 1. Performance of Detectors

그림 1에 여러 검파기의 검파 확률의 한 보기를 보였다. 이 그림의 점들은 오경보확률(fals-

e alarm probability)이 $P_{fa}=10^{-3}$ 일 때 상대 오차가 0.01 씬이 되도록 하려고 100,000번 반복 실행하여 얻었다. 한편 국소 최적 순위 검파기의 검파 확률을 얻을 때에는 $c_2(i,k)$ 를 $c_1(i)c_1(k)$ 로 근사화한 검정 통계량을 썼다. 이 그림으로부터 국소 최적 순위 검파기는 국소 최적 검파기보다 성능이 조금 나쁘지만 다른 검파기 보다는 성능이 나음을 알 수 있으며, 이것은 점근 성능 비교와 거의 비슷한 결과임을 보여준다.

V. 맺음말

이 논문에서는 가산성 잡음이 섞여 있을 때 확률 신호를 국소 최적 순위 검파하는 것에 대하여 알아 보았다. 확률 신호를 국소 최적 순위 검파할 때 쓰이는 검정 통계량이 확률 신호를 국소 최적 검파할 때 쓰이는 검정 통계량과 비슷함과 알려진 신호를 국소 최적 순위 검파할 때 쓰이는 검정 통계량을 일반화한 것임을 보였다. 이와 더불어 확률 신호를 국소 최적 순위 검파하는 검파기의 점근 성능과 유한 표본 성능을 다른 검파기들과 견주어 보았다. 국소 최적 순위 검파기는 국소 최적 검파기 보다는 성능이 조금 떨어지지만 다른 검파기 보다는 성능이 나음을 알 수 있었다. 이 논문의 결과들은 수중신호 검파, 지진파 검파 문제를 포함하는 많은 실질적인 분야에 쓰일 수 있을 것이다.

감사의 글

이 논문을 심사하신 분께 고마움을 나타냅니다.

참고 문헌

1. S.A. Kassam, *Detection of Signals in Non Gaussian Noise*, Springer Verlag, New York, 1988.
2. S.A. Kassam and J.B. Thomas, *Nonparametric Detection Theory and Applications*, Dowden, Hutchinson & Ross, Inc., Stroudsburg, 1980.
3. I. Song and S.A. Kassam, "Locally Optimum Detection of Signals in a Generalized Observation Model: The Random Signal Case", *IEEE Trans. Inform. Theory*, Vol. IT-36, pp.516-530, May 1990.
4. I. Song, J.C. Son and K.Y. Lee, "Detection of Composite of Composite Signals: Part I, Locally Optimum Detector Test Statistics", *Signal Processing*, vol.22, pp. , April 1991 (to be published).
5. J. Hajek and Z. Sidak, *Theory of Rank Tests*, Academic, New York, 1967.
6. E.L. Lehmann, *Testing Statistical Hypotheses*, 2nd ed., John Wiley & Sons, New York, 1981.
7. T. S. Ferguson, *Mathematical Statistics: A Decision Theoretic Approach*, Academic, New York, 1967.
8. I. Song, *A Nonparametric Scheme for Detection of Random Signal*, Final Report, KAIST, Seoul, February 1989.

부록.검정 통계량 유도

먼저 다음과 같이 정의하자:

$$V(x, \theta) \triangleq \prod_{i=1}^n f_{s_i}(x_i - \theta s_i) \tag{A.1}$$

$$U_i(\theta) \triangleq g(x_i - \theta s_i). \tag{A.2}$$

그러면 $g(x) = \text{sgn}(x)g(|x|)$, $f_{s_i}(-x) = f_{s_i}(x)$ 및 $h(|x|) = h(x)$ 임을 쓰면 다음을 얻을 수 있다.

$$\frac{d^2 P(q, z | K)}{d\theta^2} = \int_B \sum_{i=1}^n \int_{K^n} s_i \frac{d}{d\theta} \{V(x, \theta) U_i(\theta)\} f(s) ds dx, \tag{A.3}$$

(A.3)에서 $B = \{x | Q=q, Z=z\}$ 이다. 이제

$$U_i(0) = g(x_i), \tag{A.4}$$

$$V(x, 0) = f(x | H), \tag{A.5}$$

$$U_i|_{\theta=0} = s_i \{h(x_i - g^2(x_i))\} \tag{A.6}$$

및

$$\begin{aligned} V(x, \theta)|_{\theta=0} &= \{V(x, \theta) \sum_{k=1}^n s_k U_k(\theta)\}|_{\theta=0} \\ &= V(x, 0) \sum_{k=1}^n s_k g(x_k) \end{aligned} \tag{A.7}$$

임을 염두에 두면 다음을 얻을 수 있다.

$$\begin{aligned} & \frac{d^2 P(q, z | K)}{d\theta^2} |_{\theta=0} \\ &= \int_B \sum_{i=1}^n \int_{K^n} s_i [V(x, 0) s_i \{h(x_i - g^2(x_i))\} + V(x, 0) \sum_{k=1}^n s_k g(x_k) g(x_i)] f(s) ds dx \\ &= \sum_{i=1}^n \sigma_i^2 \int_B V(x, 0) \{h(x_i) - g^2(x_i)\} dx + \sum_{i=1}^n \sum_{k=1}^n \int_{K^n} s_i s_k f(s) ds \int_B V(x, 0) g(x_k) g(x_i) dx \\ &= \sum_{i=1}^n \sigma_i^2 \int_B V(x, 0) h(|x_i|) dx + \sum_{i=1}^n \sum_{k=1, k \neq i}^n \int_{K^n} s_i s_k f(s) ds \int_B z_i z_k V(x, 0) g(x_k) g(x_i) dx \\ &= [\sum_{i=1}^n \sum_{k=1, k \neq i}^n r_{s_i}(i, k) z_i z_k c_2(q, q_k) + \sum_{i=1}^n \sigma_i^2 d(q_i)] / (2^n n!) \end{aligned} \tag{A.8}$$

(A.8)을 $P(q, z | H) = \frac{1}{2^n n!}$ 로 나누면 (13)을 얻는다.



宋 翊 鎬(Ick Ho SONG) 正會員
1960년 2월 20일생
1982년 2월 : 서울대학교 전자공학과 졸업
1984년 2월 : 서울대학교 전자공학과 졸업(석사)
1985년 8월 : Univ. of Pennsylvania 전기 공학과 졸업(M.S.E)
1987년 5월 : Univ. of Pennsylvania 전기 공학과 졸업(Ph. D.)

현재 : 한국과학기술원 전기 및 전자공학과 조교수



吳 澤 相(Taek Sang OH) 正會員
1963년 10월 28일생
1988년 2월 : 인하대학교 전자공학과 졸업
1981년 2월 : 한국과학기술원 전기 및 전자공학과 졸업(석사)
현재 : 삼성사 중앙연구소 연구원



嚴 泰 相(Tae Sang UHM) 正會員
1965년 2월 7일생
1988년 2월 : 연세대학교 전자공학과 졸업
현재 : 한국과학기술원 전기 및 전자공학과 석사과정 재학



韓 永 玉(Young Ok HAN) 正會員
1966년 4월 21일생
1989년 2월 : 서울대학교 전기공학과 졸업
1991년 2월 : 한국과학기술원 전기 및 전자공학과 졸업(석사)
현재 : 삼성전자(반도체 부문) 연구원