Combining Statistical Information in Set Theoretic Estimation

Patrick L. Combettes and Thierry J. Chaussalet

Abstract— **Statistical information has been used extensively to define feasible solutions in set theoretic signal processing. However, the reliability of the set theoretic estimates resulting from the combination of multiple statistical constraints has thus far not been investigated. In this letter, we address this question in order to provide a basis for a better use of statistical information in set theoretic estimation problems.**

I. Problem Statement

In set theoretic estimation, the objective is to produce solutions that satisfy all the constraints $(\Psi_i)_{i\in I}$ available for the problem. Each constraint Ψ_i is associated with the property set $S_i = \{a \in \Xi \mid a \text{ satisfies } \Psi_i\}$ in the solution space Ξ, and the set of solutions is the feasibility set $S = \bigcap_{i\in I} S_i$ (the reader is referred to [1] for a detailed overview of the field and of its applications in signal and image processing).

In the family of constraints $(\Psi_i)_{i\in I}$, one can distinguish the statistical ones $(\Psi_j)_{j\in J\subset I}$ from the deterministic ones $(\Psi_i)_{i\in I\setminus J}$. Statistical constraints typically arise from known properties (e.g., bounds, moments, mixed moments, dependence, spectral and polyspectral properties) of the random processes involved in the data generation model. They have been used at various levels of generality in a number of studies, e.g., [1]–[4], [7]–[9]. Since a constraint Ψ_j is enforced via a statistical test, the resulting property set S_j is actually a confidence region that contains the true object being estimated h to within some confidence level p_j. To evaluate the reliability of a set theoretic estimation procedure, the relevant figure is the confidence level of the feasibility set, i.e., the probability c that S contains h. The prevalent approach, however, has been to choose *ad hoc* values for the coefficients $(\mathsf{p}_j)_{j\in J}$ without paying much attention to the problem of securing a reasonable global confidence level c. Consequently, the feasibility set thus obtained may not contain h and may even be empty. As an example, consider the image processing problems of [2], [7], and [8], where a statistical constraint was associated with each of the N^2 pixels in the image. Then, since the constraints are independent, the standard confidence level $\mathsf{p}_j = 99\%$ on each set makes the reliability of the resulting set theoretic estimates collapse to $\mathsf{c} = \mathsf{p}_j^{N^2} \approx 0$. This example underlines the need for a statistical analysis of the reliability of set theoretic estimates in order to lay the groundwork for the rational and productive use of statistical constraints in set theoretic estimation.

Manuscript received May 19, 1995. The associate editor coordinating the review of this letter and approving it for publication was Dr. R. Ansari.

P. L. Combettes is with the Department of Electrical Engineering, City College and Graduate School, City University of New York, New York, NY 10031, USA.

T. J. Chaussalet is with the Division of Mathematical and Decision Sciences, School of Computer Science and Information Systems Engineering, University of Westminster, London, UK.

Publisher Item Identifier S 1070-9908(96)02745-9.

II. Statistical Analysis

The object to be estimated is denoted by h and belongs to a vector space Ξ. The observable data process $(X_n)_{n\in\mathbb{Z}}$ is modeled as

$$(\forall n \in \mathbb{Z})(\forall \omega \in \Omega)\ \ X_n(\omega) = T_n(h) + U_n(\omega) \tag{1}$$

where we define

$(\Omega, \mathcal{F}, \mathsf{P})$	underlying probability space
$(T_n)_{n\in\mathbb{Z}}$	known sequence of functionals on Ξ modeling the underlying signal formation process
$(U_n)_{n\in\mathbb{Z}}$	random sequence representing the uncertainty, e.g., measurement noise and/or inaccuracies in the signal formation model.

Now, let Ψ_j $(j \in J)$ be a known property of the uncertainty process $(U_n)_{n\in\mathbb{Z}}$. For every estimate $a \in \Xi$ of h, we define a residual process $(Y_n(a))_{n\in\mathbb{Z}} = (X_n - T_n(a))_{n\in\mathbb{Z}}$. Since, theoretically, $(Y_n(h))_{n\in\mathbb{Z}} = (U_n)_{n\in\mathbb{Z}}$, we construct the property set

$$S_j = \{a \in \Xi \mid (Y_n(a))_{n\in\mathbb{Z}} \text{ satisfies } \Psi_j\}. \tag{2}$$

Such a set cannot be used directly since only a finite trajectory $Y(a,\omega) = (X_n(\omega) - T_n(a))_{1\le n\le N}$ of the residual process is observed for a realization $\omega \in \Omega$. Let $Q_j(a)$ be a statistic of $(X_n - T_n(a))_{1\le n\le N}$ describing the property Ψ_j, and let p_j be its desired confidence coefficient, that is

$$\mathsf{p}_j = \mathsf{P}\{\omega \in \Omega \mid h \in S_j(\omega)\}. \tag{3}$$

Then, from the distribution of $Q_j(h)$, we can find a Borel set $\mathcal{R}_j$ such that $\mathsf{P}\{\omega \in \Omega \mid Q_j(h,\omega) \in \mathcal{R}_j\} = \mathsf{p}_j$. We thus replace the exact property "$(Y_n(h))_{n\in\mathbb{Z}}$ satisfies Ψ_j" with the property "$Q_j(h,\omega) \in \mathcal{R}_j$." It expresses the fact that h generates a residual trajectory $Y(h,\omega)$, which agrees with Ψ_j at a confidence level p_j. This yields the set

$$S_j(\omega) = \{a \in \Xi \mid Q_j(a,\omega) \in \mathcal{R}_j\}. \tag{4}$$

Let us define the events

$$(\forall j \in J)\ \ E_j = \{\omega \in \Omega \mid Q_j(h,\omega) \in \mathcal{R}_j\} \tag{5}$$

$$= \{\omega \in \Omega \mid h \in S_j(\omega)\} \tag{6}$$

1070–9908/96$05.00 © 1996 IEEE

and

$$E = \{\omega \in \Omega \mid h \in S(\omega)\} = \bigcap_{j \in J} E_j. \tag{7}$$

Then, $(\forall j \in J)\ \mathsf{P}E_j = \mathsf{p}_j$ and $\mathsf{P}E = \mathsf{c}$. The problem is to determine the confidence levels $(\mathsf{p}_j)_{j\in J}$ to be placed on the sets $(S_j)_{j\in J}$ to achieve a preset confidence level c on the solution set S.

Theoretically, if F designates the joint distribution function of the statistics $(Q_j)_{j\in J}$, then the global confidence level can be expressed as

$$\mathsf{c} = \int_{\mathbf{X}_{j\in J}\mathcal{R}_j} dF. \tag{8}$$

In practice, F is rarely tractable, and this theoretical expression is difficult to evaluate. However, the joint distribution functions of subgroups of statistics are usually known exactly or asymptotically [3], [5]. In order to derive a lower bound on c, we shall need the following simple fact, which is known as Bonferroni's inequality [6].

Lemma: Let $(A_i)_{i\in L} \subset \mathcal{F}$ be a finite collection of events. Then

$$\mathsf{P}\bigcap_{i\in L} A_i \geq 1 - \sum_{i\in L}(1 - \mathsf{P}A_i). \tag{9}$$

As noted in [6], this bound is quite useful and surprisingly tight when the probabilities $(\mathsf{P}A_i)_{i\in L}$ are close to 1. Let us also note that the product $\prod_{i\in L}\mathsf{P}A_i$ majorizes Bonferroni's bound by virtue of the inequality $\prod_{i\in L}(1-\alpha_i) \geq 1-\sum_{i\in L}\alpha_i$, for $(\alpha_i)_{i\in L} \subset]0,1[$; therefore, it cannot be used to bound c unless, of course, the events are independent.

Now, let $(J_l)_{l\in K}$ be a partition of J such that the probabilities

$$(\forall l \in K)\ \ \mathsf{p}_l = \mathsf{P}\bigcap_{j\in J_l} E_j \tag{10}$$

are computable via the (exact or asymptotic) distribution functions $(F_l)_{l\in K}$ of the statistics $((Q_j)_{j\in J_l})_{l\in K}$. If we let $(\forall l \in K)\ \ F_l = \bigcap_{j\in J_l} E_j$, then $E = \bigcap_{l\in K} F_l$, and (9) yields

$$\mathsf{c} = \mathsf{P}\bigcap_{l\in K} F_l \geq 1 - \sum_{l\in K}(1 - \mathsf{p}_l). \tag{11}$$

This inequality shows how the probabilities $(\mathsf{p}_l)_{l\in K}$ should be chosen in order to secure a given global confidence level c. For instance, when the probabilities $(\mathsf{p}_l)_{l\in K}$ are identical and equal to p, we derive from (11) that

$$\mathsf{p} = 1 - (1 - \mathsf{c})/k \tag{12}$$

where $k = \mathrm{card}(K)$. Furthermore if, for every $l \in K$, each subgroup $(Q_j)_{j\in J_l}$ consists of k_l independent statistics with equal confidence levels, it follows from (12) that

$$(\forall l \in K)(\forall j \in J_l)\ \ \mathsf{p}_j = (1 - (1 - \mathsf{c})/k)^{1/k_l}. \tag{13}$$

III. Application to Signal Restoration

To illustrate an application of the general analysis of the previous section, we now consider the following set theoretic signal restoration problem [2], [8]. An original N-point signal h is to be estimated from an N-point degraded signal $x = Th + u$, where the $N \times N$ matrix T represents a known convolutional blur, and where u is an N-point noise vector whose components are drawn from a zero-mean Gaussian white population $(U_n)_{n\in\mathbb{Z}}$ with $\mathsf{E}|U_0|^2 = \sigma^2$. From this information, the following sets can be constructed, where T_n designates the nth row of T and $\mathrm{erf}(x) = (2/\sqrt{\pi})\int_0^x e^{-t^2}dt$ [3], [8].

- N sets arising from range information: $S_n^R = \{a \in \mathbb{R}^N \mid Q_n^R(a) \in \mathcal{R}^R\}$ $(1 \leq n \leq N)$, where

$$\begin{cases} Q_n^R(a) = |x_n - T_n{\cdot}a| \\ \mathcal{R}^R = [0, \sqrt{2}\sigma\mathrm{erf}^{-1}(\mathsf{p}_R)]. \end{cases} \tag{14}$$

- A set arising from moment information: $S^V = \{a \in \mathbb{R}^N \mid Q_V(a) \in \mathcal{R}_V\}$, where

$$\begin{cases} Q^V(a) = \left|\sum_{n=1}^N |x_n - T_n{\cdot}a|^2 - N\sigma^2\right| \\ \mathcal{R}^V = [0, 2\sigma^2\sqrt{N}\mathrm{erf}^{-1}(\mathsf{p}_V)]. \end{cases} \tag{15}$$

- $N/2 - 1$ sets arising from spectral information: $S_k^P = \{a \in \mathbb{R}^N \mid Q_k^P(a) \in \mathcal{R}^P\}$ $(1 \leq k \leq N/2 - 1)$, where

$$\begin{cases} Q_k^P(a) = \frac{2}{N}\left|\sum_{n=1}^N (x_n - T_n{\cdot}a)\exp(-i\frac{2\pi}{N}kn)\right|^2 \\ \mathcal{R}_P = [0, -N\sigma^2\ln(1 - \mathsf{p}_P)]. \end{cases} \tag{16}$$

Now, suppose that a desired global confidence coefficient is $\mathsf{c} = 95\%$ and, for sake of definiteness, take $N = 64$. The confidence coefficients p_R, p_V, and p_P must be determined accordingly. Three groups of sets will be considered: $S^R = \bigcap_{n=1}^N S_n^R$, S^V, and $S^P = \bigcap_{k=1}^{N/2-1} S_k^P$. We shall call p the confidence level on S^R, S^V, and S^P. From (12), we get that $\mathsf{p} = 1 - (1 - \mathsf{c})/3 = 98.333\%$. Note that $\mathsf{p}_V = \mathsf{p} = 98.333\%$. Next, since the statistics $(Q_n^R)_{1\leq n\leq N}$ are independent, we get $\mathsf{p}_R = \mathsf{p}^{1/N} = 99.974\%$. Likewise, since the statistics $(Q_k^P)_{1\leq k\leq N/2-1}$ are independent, (13) yields $\mathsf{p}_P = \mathsf{p}^{1/(N/2-1)} = 99.946\%$. Let us note that enforcing the standard 99% confidence level on each of the sets $(S_n^R)_{1\leq n\leq N}$, S^V, and $(S_k^P)_{1\leq k\leq N/2}$ would guarantee a global confidence level of only $c' \geq 0.99^N + 0.99 + 0.99^{N/2-1} - 2 = 24.790\%$.

References

[1] P. L. Combettes, "The foundations of set theoretic estimation," *Proc. IEEE*, vol. 81, no. 2, pp. 182–208, Feb. 1993.

[2] P. L. Combettes and H. J. Trussell, "Methods for digital restoration of signals degraded by a stochastic impulse response," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 3, pp. 393–401, Mar. 1989.

[3] ——, "The use of noise properties in set theoretic estimation," *IEEE Trans. Signal Processing*, vol. 39, no. 7, pp. 1630–1641, July 1991.

[4] J. R. Deller, Jr., M. Nayeri, and S. F. Odeh, "Least-square identification with error bounds for real-time signal processing and control," *Proc. IEEE*, vol. 81, no. 6, pp. 815–849, June 1993.

[5] M. Fisz, *Probability Theory and Mathematical Statistics*, 3rd ed. New York: Wiley, 1963.

[6] R. G. Miller, Jr., *Simultaneous Statistical Inference*, 2nd ed. New York: Springer-Verlag, 1981.

[7] M. K. Özkan, A. M. Tekalp, and M. I. Sezan, "POCS-based restoration of space-varying blurred images," *IEEE Trans. Image Processing*, vol. 3, no. 4, pp. 450–454, July 1994.

[8] H. J. Trussell and M. R. Civanlar, "The feasible solution in signal restoration," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-32, no. 2, pp. 201–212, Apr. 1984.

[9] E. Walter (Guest Editor), "Parameter identifications with error bounds," *Math. Comput. Simul.*, vol. 32, no. 5–6, Dec. 1990.