ROBUST FILTERING AND FAULT DIAGNOSIS BY GNOSTICAL METHODS

P. Kovanic, J. Pacovský
Institute of Information Theory and Automation
Czechoslovak Academy of Sciences
Pod vodárenskou věží 4, 182 08 Praha 8

Abstract. Some practical applications of methods based on the gnostical theory are briefly described. Gnostics is an alternative to statistics. It develops methods not based on statistical models. These methods may be applied also to small samples of bad data and in nonstationary situations. Examples demonstrate nonlinear filters of gnostical type, which are

- robust with respect to disturbances of data
- sensitive to actual changes of the process represented by data
- adaptive in all actions
- capable to derive reliable information on sudden changes of data

Such a high intelligency of gnostical algorithms makes them suitable especially for supervision systems, alarm monitoring, fault detection and diagnosis, quality control.

Keywords. Quality control; adaptive systems; filtering; automatic testing; fault detection and diagnosis; gnostics.

INTRODUCTION

A new theory applicable to broad scale of control problems including estimation and identification , prediction and filtering was developed (Kovanic 1984 a, b, c) and briefly exposed at IFAC '85 Congress (Kovanic 1985). A new exposition of the theory with examples of applications can be found in (Kovanic 1986). This ("gnostical") theory does not use any statistical assumptions on uncertainty of data. It is based on a couple of realistic axioms. The first axiom supposes the data to be disturbed either only additively or only multiplicatively. This axiom was shown in Kovanic (1984 a) to determine fully the following:

- the metric (of a non - Euclidean, Riemannian type) of errors of uncertain data
- the formulae for evaluation of the information loss and entropy increase of each individual datum due to uncertainty
- the ideal way of estimation of the true value of a datum, which minimizes the necessary information loss and entropy increase.

The second axiom of the gnostical theory is the data composition law determining the way of using a data sample for optimal estimation of the common true value of the data. Motivation of this axiom is the aim to ensure consistency between the theory of uncertain data and relativistic physics. Such a requirement is not strange for everybody who is aware of the motivation of the data composition law of the classical statistics: the arithmetical mean, variance and covariance matrix are statistical analogies of the center of masses and of the components of the inerciality tensor of a system of mass points in classical mechanics.

As shown in Kovanic (1984 b), the two axioms result in an inherent robustness of gnostical estimates. It is the aim of this paper to show by examples, that the theoretical expectation of good applicability of the gnostical theory can be supported by results of already working algorithms.

THEORETICAL BACKGROUND

An i-th datum from a data sample

$$Z(z_o,s,n) := \{z_1,\dots,z_n\}$$

is assumed to have the form

$$z_i=z_o\exp(s\,\Omega_i) \qquad (z_o,s\in R_+,\ \Omega_i\in R_1) \qquad (1)$$

where R_1 is the set of all real numbers,

R_+ is the set of all positive reals,

z_o is a "true" value of all data of the sample Z,

Ω_i represents the effect of the uncertainty of the i-th observation

s is a scale parameter characterizing the spread of data of the sample Z

The data model (1) can be easily transformed into the ordinary additive form by logarithmization. The error of the datum z_i was proved in Kovanic (1984 a) to be given by the nonlinear function

$$h(g_i) = (g_i^{-1} - g_i)/(g_i^{-1} + g_i) \quad (2)$$

where

$$g_i = (z_i/z_o)^{2/s} \quad (3)$$

A direct application of (2) is a gnostical estimate of the probability, i.e. of the quantity

$$P\ (z_o \leqslant \tilde{z}_o / Z(z_o, s, n))$$

Its gnostical estimate has the form

$$P = (1 + (\sum_i^n h\ (\tilde{g}_i))/n)/2 \quad (4)$$

where $h(\tilde{g}_i)$ is (2) with the argument g_i into which an estimate $\tilde{z}_o$ has been substituted instead of unknown true value z_o. Another application of (2) is the gnostical variance

$$V^2 = (\sum_i^n h^2(\tilde{g}_i))/n \quad (5)$$

which is closely connected with the data density of the sample $Z\ (z_o, s, n)$:

$$dP/d(\ln \tilde{z}_o) = (1 - V^2)/s \quad (6)$$

(If we would use $\tilde{z}_o$ instead of $\ln \tilde{z}_o$ in (6), we would have the estimate $dP/d\tilde{z}_o$ of the probability density of the sample.) The quantities h (2) and V^2 (5) play roles analogical as mean and variance in classical statistics. For data disturbed only slightly the gnostical error function h turns to be proportional to arithmetical mean and the gnostical variance V^2 to ordinary variance. But for data with gross errors their behaviour is quite different. As seen from the formulae the influence of the outliers (both $z_i \to \infty$ and $z_i \to 0$) is bounded in the case of gnostical quantities.

It follows from (6), that an estimate $\tilde{z}_o$ of the true value z_o, which minimizes the gnostical variance, corresponds to the location of the maximum of the data density. This estimate can be obviously obtained by solving the equation

$$d^2P/d(\ln \tilde{z}_o)^2 = 0 \quad (7)$$

with respect to the variable $\tilde{z}_o$. This nonlinear equation has a solution always, but it may have more than one solution in the case of more than one single "cluster" of data of the sample Z. If we repeat the solution process using different starting points, we can determine subsequently the locations af all maxima of the data density function.

To apply all mentioned formulae, we need an estimate $\tilde{s}$ of the scale parameter. For many practical purposes such estimate $\tilde{s}$ gave good results, which was obtained by solving the equation

$$\varphi^{-1} \sin \varphi = n^{-1} \sum_i^n \sqrt{1 - h^2(g_i)} \quad (8)$$

where

$$\varphi = \pi \tilde{s}/2 \quad (9)$$

Equation (8) can be interpreted in the framework of the gnostical theory as a condition of equivalence of the entropy in both discrete and continuous models of data uncertainty. Another approach to scale estimation is exposed in (Kovanic 1986): the "best" value of the scale parameter should minimize the maximum distance between the empirical distribution function of the data sample and the gnostical distribution function. Such scale parameter is useful especially for important applications of estimates of values of the probability distribution or of quantiles of this function. Unfortunately, this scale estimating is connected with extremalization of a non-smooth function and requires tedious calculations. For many on-line applications such as filtering a fast algorithm solving (8) is a good compromise between quality of results and calculation effort.

It is interesting to rewrite (7) into a more readable form of the equation

$$\tilde{z}_o = ((\sum_i^n f_i^3\ z_i^{2/s})/(\sum_i^n f_i^3\ z_i^{-2/s}))^{s/4} \quad (10)$$

where

$$f_i = 2/((z_i/z_o)^{2/s} + (\tilde{z}_o/z_i)^{2/s}) \quad (11)$$

If all z_i deviate from $\tilde{z}_o$ only slightly (the case of very small errors), then the estimate $\tilde{z}_o$ satisfying the (10) approaches the arithmetical mean $\sum_1^n z_i\ /n$. Quite different is the case of strong errors: the weight f_i (called the fidelity in gnostical theory) decreases rapidly for values of ratio $z_i/\tilde{z}_o$ deviating from 1. The outlying data have thus a negligible effect on the solution $\tilde{z}_o$ of (10). This is why the estimate $\tilde{z}_o$ is very robust with respect to outliers.

ROBUST FILTER OF A GNOSTICAL TYPE

We are ready to consider now some features of gnostical algorithms using practical examples. There is a real process depicted in Fig. 1 by points as a time series Z(t) . The quantity Z is an important parameter of quality of a chemical product, the % of water content. The gnostical filter will be an algorithm treating an n-tuple of last observed values of the time series as a data sample Z to find an estimate $\tilde{z}_o$ of its location parameter z_o as a solution of the equation (7) using the estimate $\tilde{s}$ of the scale parameter s obtained from (8). Then the sample Z is actualized by substituting a new observed value of the series instead of the eldest one. The estimate $\tilde{z}_o$ is thus the output of the filter being actualized after each step of observation. The full line in Fig. 1 connects the output values of such a filter. The size n of the treated sample Z equals to 10. This means that the time response of the filter does not exceed time of 10 observations.

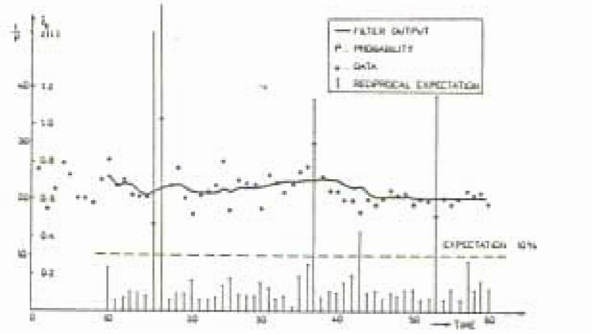

Fig. 1. Robust filtering of a time series and reciprocal expectation of the newest datum

The high robustness of the filter manifests e. g. in Fig. 1 by its insensitivity to outlying observations at time 15, 16, 37. The robustness of the gnostical filter is not paid by a slow response with respect to actual changes of process level. To show this, the same data as in Fig. 1 have been used and modified starting at time 32 by multiplication by 1.2 to produce a step function. As seen in Fig.2, the filter needs only 7 observations to go over to the new process level.

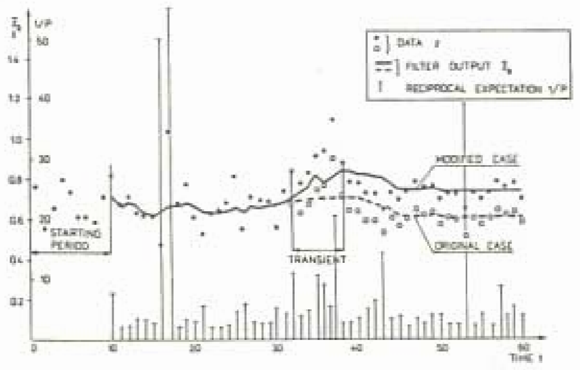

Fig. 2. Step response of a gnostical filter

High robustness of the solution of (10) can be demonstrated by another examples. Data sample Z_A in Fig.3 consists of actual values of some real data. Sample Z_C is obtained from the sample Z_A by multiplication of all data by 1.2. "Mixed" sample Z_B differs from the sample Z_A only by the first datum which is increased by 20%. The last sample Z_D differs from the increased data of the sample Z_C by the last datum which has the original value. Gnostical scale parametr $\tilde{S}$ estimated from the samples Z_A and Z_C is $\tilde{S}_A = \tilde{S}_C = 0.06338$. Let us use an algorithm solving the equation (10) with respect to $\tilde{Z}_0$ to find the location parameters of the samples. Results are summarized in Table 1 together with arithmetical means $\bar{z}$ of samples.

TABLE 1. Estimated location parameters (example in Fig. 3)

Data sample	Arith. mean $\bar{z}$	Gnostical parameter of location	
		$\tilde{Z}_{o1}$	$\tilde{Z}_{o2}$
Z_A	9.468	9.398	–
Z_B	9.654	9.420	11.158
Z_C	11.362	–	11.277
Z_D	11.714	9.5105	11.250

For "clean" samples Z_A and Z_C the gnostical location parameters differ only slightly from arithmetical means. Small change of one single datum of Z_A causes a change of the arithmetical mean by 2% while the gnostical location parameter increases only by about 0.2%. Another important feature is that there are two solutions of the equation (10) in the case Z_B . The second parameter $\tilde{Z}_{o2}$ locates nearly exactly at the changed value of the first datum of Z_B . Similar situation occurs in the case of Z_D. But let us compare the locations for both pairs of samples. It is not surprising that the ratio $\tilde{Z}_{o2C}/\tilde{Z}_{o1A}$ of location parameters of clean samples equals to 1.2 because of proportionality of all data. However, it is not trivial that the ratio $\tilde{Z}_{o2D}/\tilde{Z}_{o1B}$ is 1.194. These estimates were obtained by nonlinear operations applied to nonproportional samples and therefore such a small change of the ratio of location parameters speaks on the robustness of estimates.

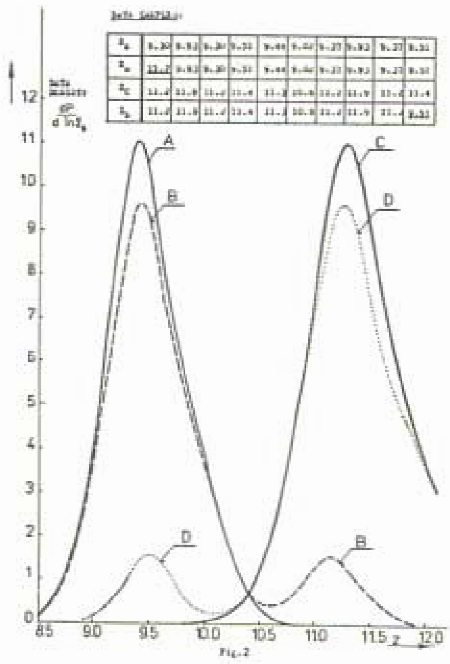

Fig. 3. Four examples of data samples and their data density curves

On the other hand, the appearance of second location parameters resulting from the slight modification of a single datum demonstrates high resolution power of the method. It can be explained simply by having a look at Fig. 3 showing data densities of all four samples.

These examples demonstrate a quite unusuall feature of the gnostical filter: It can be applied to data having a bimodal (or even multimodal) density function. In the bimodal case it is sufficient to start the iterations solving the equation (11) twice - from the minimal and maximal data values. Such a filter has two outputs characterizing the levels (the locations of maximal density) of both "mixed" processes. It can be also used as an indicator of the "homogenity" of the process: A process is homogenious (its density is unimodal) when both outputs coincide.

DIAGNOSTIC FUNCTIONS OF THE GNOSTICAL FILTER

Expectation of a Datum

Having some n last observations z_i of a process we can make use of (4) to estimate the probability P (the "expectation") of an arbitrary value of a new datum. We may thus confront the newly obtained measurement with its expectation to guess in which degree it may be a "proper" continuation of the previous n values. If the expectation appears to be too low, a diagnostic signal may be automatically initiated. To show this, the values of reciprocal expectations are depicted by vertical bars in Fig.1. The horizontal dotted line corresponds to the reciprocal expectation 1/0.1 which is expected to be exceeded only in 10% cases. As shown in Fig.1, the reciprocal expectation reliably marks the outlying data. It is important, that this function of the gnostical algorithm is fully adaptive with respect to changes of both process level and intensity of disturbances. This is demonstrated by Fig. 2 where the reciprocal expectation is a useful information even during a transient.

Automatic Classification of Data

This function deserves to be described in more details. We consider again a part of a real data series to form a data sample Z having ten elements. Nine of them are 8.95, 9.44, 9.30, 9.86, 9.30, 9.58, 9.44, 10.14, 9.09 and the tenth datum z_{10} will be taken as a variable going through the whole interval from 0 to infinity. We estimate the parameter of location of the sample Z for various values of z_{10} to show its sensitivity with respect to values of z_{10} given $z_1, \dots, z_9$. This sensitivity is demonstrated by Fig.4. We see that one single value can never change the location parameter of the sample more than about ± 0.6%. The locations z_L and z_U of the maximum and minimum of the sensitivity curve are taken to be the bounds of "typical" values of the data, while the values $\tilde{Z}_o(z_L)$ and $\tilde{Z}_o(z_U)$ are bounds of the

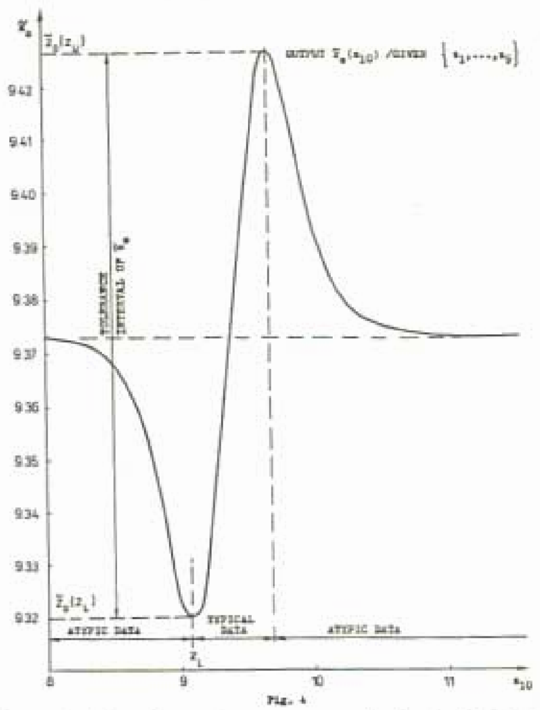

Fig. 4. The dependence z_o of the location parameter of a sample $z_1, \dots, z_{10}$ on the changing value of the z_{10}(given fixed $z_1, \dots, z_9$)

"tolerance interval" of the location parameter. Data Z may be now classified as typical ($z_L \leq z \leq z_U$), subtypical ($z < z_L$) and supertypical ($z_U < z$). The ratios of numbers of data of different classes may be used as diagnostic indicators of stationarity of the observed process. The behaviour of the bounds of typical data can be demonstrated in Fig. 5 using the same data as in Fig. 1.

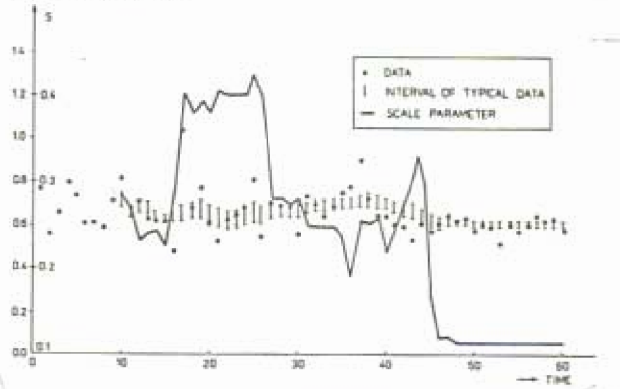

Fig. 5. Intervals of typical data and scale parameter of the same data series as in Fig. 1

The evaluation of these bounds can be easily performed using relations

$$z_L = \tilde{Z}_o/K \qquad z_U = \tilde{Z}_o/K \qquad (12)$$

where

$$K = ((\sqrt{3} + 1)/\sqrt{2})^{s/2} \qquad (13)$$

Proof of (12) and (13) will be published elsewhere.

The stabilization of the process at right hand part of Fig. 5 results in a narrowing of the interval of typical data because of the decreased scale parameter $\tilde{s}$.

Scale Parameter as a Diagnostic Indicator

There is also the scale parameter $\tilde{s}$ shown by the full line in Fig. 5. It characterizes reliably the spread of the data as well as the appearance of individual outliers and sharp changes of the observed process as a whole. It may be therefore also used as an indicator of undesirable changes of the observed data series for diagnostic purposes.

THE IDEA OF A GNOSTICAL MONITOR

We have shown above that gnostical algorithms can be used not only as efficient robust filters but also as sensitive and reliable diagnostic means. A natural idea appears to integrate all such functions into a new type of real-time program. Such a program is called the gnostical monitor. It has been already developed and applied in industry for supervision of processes, in emergency systems and in a technological control system for quality control of production.

OTHER TYPES OF GNOSTICAL MONITORS

The gnostical monitor working according to this paper has been marked as GM1. Another one, GM2 uses another gnostical formulae for the filtering procedure based on gnostical identification of a linear regression model. This procedure is of a recursive type. Therefore the monitor GM2 can be used especially for real-time application requiring minimal computing time and memory. It applies exponential forgetting of old data. Its robustness with respect to outliers is even better then with the GM1 because an outlying value disturbs the output only once. Disturbances are here graduatelly forgetted unlike the case of "window" memory of the monitor GM1, where each disturbance affects the output two times.

The third type of gnostical monitors (GM3) has been developed for monitoring the probability of random events. The specific requirement of this application was the unbiasedness of the estimate ensuring the convergency of the nonlinear robust estimate to the classical linear one in a limit case. This problem has been also successfully solved by gnostical algorithms

Both monitors GM2 and GM3 perform the same diagnostic functions as the type GM1. Other gnostical monitors under development are

- a monitor of the process trend
- a monitor of a correlation coefficient of a couple of processes
- a monitor working as an on-line identifier of parameters of a process model

They all should also integrate robust filtering with diagnostic functions.

CONCLUSION

New types of real-time programs for supervision of processes, gnostical monitors based on gnostical theory of uncertain data integrates robust and efficient filtering with sensitive and reliable diagnostic functions. They are suitable for process control as well as for emergency systems.

REFERENCES

Kovanic, P. (1984 a) Gnostical Theory of Idividual Data. Problems of Control and Information Theory 13, 4, 259 - 274

Kovanic, P. (1984 b) Gnostical Theory of Small Samples of Real Data. Problems of Control and Information Theory 13, 5, 303 - 319

Kovanic, P. (1984 c) On Relations between Information and Physics. Problems of Control and Information Theory 13, 6, 383 - 399

Kovanic, P. (1985) A New Theoretical and Algorithmical Basis for Estimation, Identification and Information. IX-th World Congress IFAC (1984), Preprints, Vol. XI, 122 - 131

Kovanic, P. (1986) A New Theoretical and Algorithmical Basis for Estimation, Identification and Control. Automatica No 6. (1986)