MEASUREMENTS IN MATHEMATICAL MODELING AND DATA PROCESSING

William Moran and Barbara La Scala
University of Melbourne, Australia

Keywords: detection theory, estimation theory, signal processing, hypothesis testing

Contents

Summary

This chapter discusses the problem of extracting information from a noisy measured

signal. This problem can be decomposed into four separate subproblems: modeling the noise; estimating the noise parameters; estimating the signal; and determining when the signal is present. Two related aspects of signal processing that can be used to solve these problems are discussed – estimation theory and detection theory.

1. Introduction

Statistical signal processing is the extraction of information from signals. This somewhat cursory definition leads to two questions: what is a signal and what is information? No attempt will be made here to give a complete answer to either, instead we content ourselves with a discussion of some examples of each.

Signals in the context of this chapter are usually electronic but need not be. At the most general level, a signal is a measurement or observation. Examples of such are radio signals, images, radar signals, telephone signals, sonar signals, seismic signals, stock market prices, etc. All of these may be considered as consisting of some underlying information (which may have a random component) together with some random element (noise), which is obscuring that information. Information can be the content of a radio broadcast, telephone signal, or image, the time elapse to return of a radar or sonar pulse, etc. Stock market prices mark an interesting departure where the current price (which is observed) is known and the required information is a prediction of the price at a time in the future.

This chapter will describe two related aspects of signal processing – estimation theory and detection theory. In fact we could have treated both as different aspects of the same theory but it is customary to treat them separately and we shall not depart from that custom here. Here is an example of probably the most fundamental problem in statistical signal processing.

Example 1.A transmitter is intermittently sending out a signal – say a stream of bits (0's and 1's) of length 75. We can collect data for both the situations when the signal is sent out and when it is not. Can we use this data to find an efficient method of detecting when the signal is present?

To be more specific, in Figure 1 are three samples of a signal with noise followed in Figure 2 by three examples of noise alone.

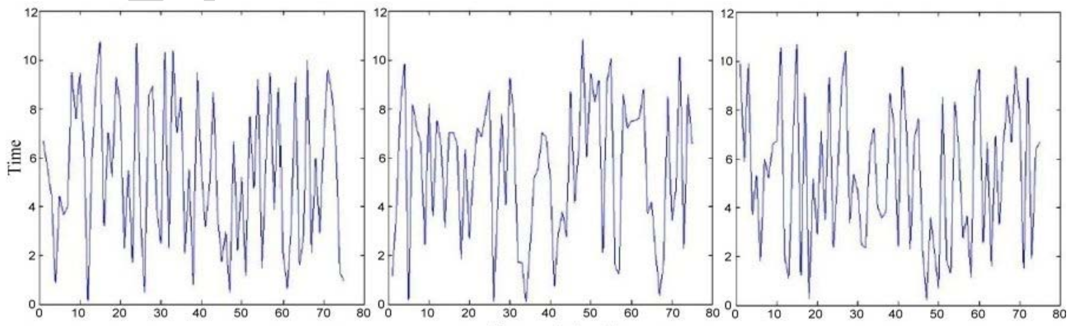

Figure 1. Signal plus noise

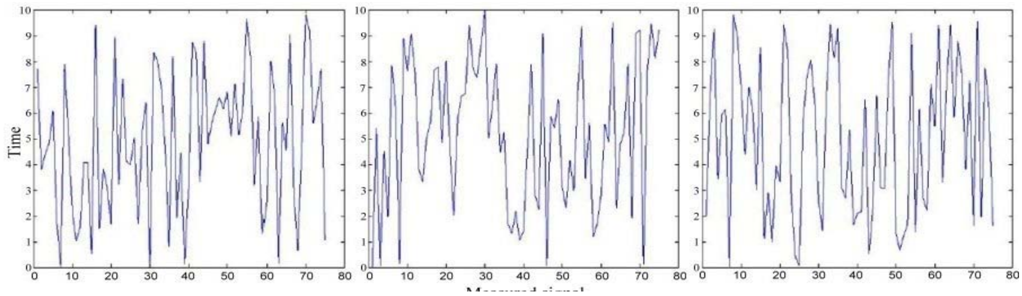

Figure 2. Noise alone

The noise is sufficient to mask any obvious visual feature of the signal, so the problem is not an easy one. The important word in the problem we have stated is “efficient”. It would not be hard to design some kind of detector, but we wish to design the “best” one. We also want to determine how well it performs. This is where the methods of statistical signal processing come in.

In fact, the above problem breaks up into four different problems:

1. Find a model for the noise.
2. Estimate the noise parameters.
3. Estimate the signal.
4. Design a detector to discriminate when signal is present and when it is not.

Our aim in this chapter is to describe and investigate methods for solving these kinds of problems.

2. Hypothesis Testing

2.1 Overview of Signal Detection

We deal first with the problem of signal detection; that is, of deciding whether the received “signal” at a receiver is merely noise or contains a specific signal, or perhaps one of a particular family of signals. The problem of signal detection is really a problem in statistical hypothesis testing. To illustrate the ideas we present a simple example.

Example 2(Binary Transmission). Binary digits (i.e., 0 or 1) are transmitted over a communication channel. Our observation of the output is one (say X) of a pair of random variables, X_0 and X_1, which are the outputs when respectively 0 and 1 are transmitted. Because of various noise and distortion problems a transmitted 0 is received as a number normally distributed with mean 0 and variance σ_0^2. Similarly, X_1 is a random variable with mean 1 and variance σ_1^2.

The concepts used in modeling this problem are essentially the same as in many more complicated ones. In this case the model is comprised of two probability distributions – the one corresponding to the received signal X_0 when a 0 is transmitted and the other to the received signal X_1 when a 1 is transmitted. The model is typically described in terms of two hypotheses to be tested:

$$\begin{aligned} H_0 &: \quad 0 \text{ was transmitted} \\ H_1 &: \quad 1 \text{ was transmitted} \end{aligned} \tag{1}$$

The probability distribution of X_0 is $\mathsf{N}(0,\sigma_0^2)$ – that is, normal with mean 0 and variance σ^2 – and that of X_1 is $\mathsf{N}(1,\sigma_1^2)$. At a later stage in the chapter we shall investigate how such information might be obtained from data.

Our aim is to take a received signal and decide whether it was associated with a transmitted 0 or 1. It should be clear that this cannot be done consistently correctly. We should expect that sometimes the noise will so distort the signal that we will choose the incorrect hypothesis. All that can be hoped to achieve is a selection mechanism which is in some sense optimal – which makes fewer mistakes than any other. The received signal will be some number, say x. A decision whether that number came from transmission of a 0 or a 1 is required.

Usually the hypotheses to be tested are asymmetric. That is, one of the hypotheses states the signal is of a different type than the other. In our context they are often

$$\begin{aligned} H_0 &: \quad \text{Signal absent} \\ H_1 &: \quad \text{Signal present} \end{aligned} \tag{2}$$

In contrast, if the two hypotheses were that the signal was normally distributed but each posited a different mean then they would be symmetric hypotheses.

For the purposes of fitting the problem we posed in terms of this model, let us regard the transmission of a 0 in the binary transmitter example as “signal absent” and transmission of a 1 as “signal present”.

In this context, notice that there are two kinds of error that can be made.

1. We believe that a signal is present when it is not.
2. We believe that a signal is absent when it is present.

These two kinds of error are, in signal processing, called, *a false alarm* and *a miss*, respectively. Other disciplines may use other terminology. In statistical books they are often called *type I errors* and *type II errors* respectively. In medical statistics they are referred to as *false positives* and *false negatives*. We shall also use the term *detection* to indicate that the signal is present and we believe that it is present.

One would like to maximize the probability of detection $\mathrm{P_d}$ (that is, minimize the miss rate) while minimizing the probability of false alarm $\mathrm{P_{fa}}$. It is always possible to achieve one of these goals, but this is generally at the expense of the other. One way to see what happens can be illustrated by plotting the detection rate against the false alarm rate. The resulting curve is usually referred to as an *ROC* curve. ROC is an abbreviation of *receiver operating characteristics*. Figure 3 is an example of a ROC curve for the simple experiment described above.

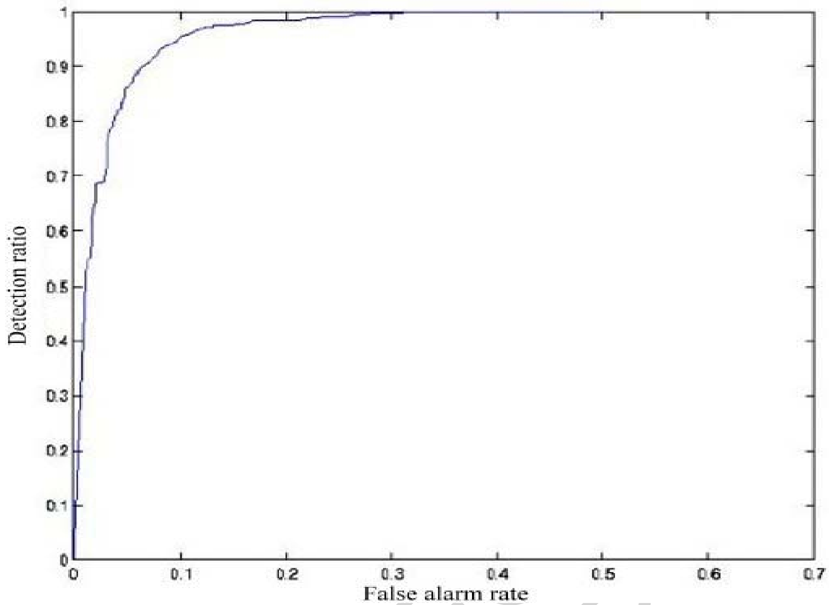

Figure 3. ROC Curve

Rather than solve Example 2 (which will be considered later) we shall seek to clarify the situation by generalizing it. In the general simple binary hypothesis test, a random vector $\mathbf{X}$ which has one of two distributions is given, thus giving rise to hypotheses:

$$\begin{aligned} H_0 &: \quad \mathbf{X} \text{ has the distribution } \mathrm{F}_0 \\ H_1 &: \quad \mathbf{X} \text{ has the distribution } \mathrm{F}_1. \end{aligned} \tag{3}$$

Write Γ for the set of possible values that $\mathbf{X}$ can take – the *observation space*. A solution of this problem will be a way of taking a member of the set of observations – one of the possible values of $\mathbf{X}$, and deciding whether we believe that value comes from distribution F_0 or distribution F_1. This is a *decision rule*. We write δ for such a decision rule. In mathematical terms the decision rule can be regarded as a partition $\delta = (\Gamma_0, \Gamma_1)$ of the set Γ. The function δ is to be chosen in some optimal way. Another (and often more convenient) way of regarding a decision rule is as a function (also called δ) from the set Γ to the set $\{0, 1\}$. Thus if $\delta(x) = 1$ for some observation x we choose hypothesis H_1, if $\delta(x) = 0$ we choose H_0.

There is a trade-off between the two kinds of errors. More false alarms in general mean fewer misses and *vice versa*. The notion of an optimal detector relies on combining the false alarm rate and miss rate into a single quantity – some kind of cost function – with respect to which the detector's performance is to be optimized. Different cost function definitions lead to different detection methodologies. The choice of detection technique depends on the situation and what is to be achieved.

The following sections consider three different kinds of cost functions and the corresponding techniques:

1. Bayesian Detection;
2. Neyman-Pearson Detection; and
3. Minimax Detection

In each case, a decision rule δ will be found to assign to any datum

$$\mathbf{x} = (x_1, x_2, \ldots, x_M)^{\mathrm{T}} \tag{4}$$

either 0 or 1, according to whether H_0 or H_1 is selected. Moreover this decision rule will be optimal with respect to some method of evaluating "cost". The different methodologies arise from different ways of assigning cost.

-
-
-

Bibliography

[1] Kay, S. M. (1998). *Fundamentals of statistical signal processing : detection theory*, volume 2. Prentice Hall. [A broad introduction to the theory and application of statistical hypothesis testing for the detection of signals in noise.]

[2] Poor, H. V. (1994). *An introduction to signal detection and estimation*. Springer-Verlag, 2nd edition. [A comprehensive and accessible introduction to detection and estimation techniques.]

[3] Scharf, L. L. (1991). *Statistical signal processing : detection, estimation and time series analysis*. Addison-Wesley.[Presents fundamental ideas in statistical signal processing in 3 areas - decision theory; estimation theory; and time series analysis.]

Biographical Sketches

Professor Moran,has been involved in many projects in and around signal processing over the last 13 years. In addition to his position at the University of Melbourne, he serves as a consulting mathematician to the Australian Defence Science and Technology Organisation (DSTO) and was also the Head of the Analytical Techniques and Medical Signal Processing Groups in the Cooperative Research Centre for Sensor Signal and Information Processing (CSSIP). Professor Moran has participated in numerous signal processing research projects for U.S. and Australian government agencies and industrial sponsors. He has published extensively in both pure and applied mathematics and is a Fellow of the Australian Academy of Science. He has authored or co-authored well over 100 published mathematical research articles.

Dr La Scala, received her BSc in Statistics in 1987 from the University of Melbourne and her PhD in Systems Engineering from the Australian National University in 1994. She is currently a member of the Melbourne Systems Laboratory (MSL) in the Electrical and Electronic Engineering Department at the University of Melbourne, Australia. Her research interests are in tracking, data fusion and network resource management. She is the Assistant Editor-in-Chief for the Journal of Advances in Information

Fusion and Associate Editor of the IEEE Transactions on Aerospace and Electronic Systems for papers on target tracking and multi-sensor systems. Before joining MSL she worked for RLM Pty Ltd designing target tracking and sensor registration algorithms for Australia's Jindalee Over-the-horizon Radar Network (JORN).