PRECISION LIMITS IN LINE PROFILE FITTING EXPERIMENTS

J. Ireland
L3Com/GSI, NASA’s GSFC, Bldg. 26, Rm. G-1, Greenbelt, MD 20771, USA

ABSTRACT

Solar spectra can provide detailed diagnostics on the state of their emitting volume. Emission line diagnostics are found by assuming a model for the spectral emission line and then fitting the model to the data. It is shown for Poisson noisy emission line data, via the application of Cramér-Rao lower bounds, that there are limits to the precision that line fitting can achieve. The limits depend on the spectral line model and the noise properties of the data. An application of the Cramér-Rao theory are presented.

1. INTRODUCTION

Solar physics, like many other sciences, makes extensive use spectrometers to measure emission lines. By modeling the way the solar physics community commonly approaches the problem of fitting an emission line model to spectral data and using Cramér-Rao lower bound theory, a framework can be constructed in which questions concerning the precision limits of idealized experiments can be answered.

Consider an experiment in which multiple measurements of the same quantity are made. Each time the experiment is run, the value observed is a sample from the distribution of measurements possible. The experiment is said to be precise if a “large proportion” of the samples lie “close” to the sample mean. In addition, the experiment is said to be accurate if the sample mean is “close” to the true value (the meaning of “close” is deliberately vague, as this depends on the purpose of the experiment). This definition of precision does not imply that a precise experiment need be accurate, as the sample mean could be very far from the true value, yet be very precisely measured. Systematic errors in the experiment setup are one reason why a precise experiment need not be accurate.

One way of measuring if a large proportion of the samples are close to the sample mean is to calculate the standard deviation of the experiment; smaller standard deviations imply that the experiment is more precise. The standard deviation is an unbiased estimator of the square root of the true variance of the distribution from which the measurements are drawn. Cramér-Rao lower bound theory (Rao 1945; Cramér 1946; Kendall & Stuart 1973) describes lower limits to the true variance and so can be interpreted as describing precision limits to experiments (Barford, 1985).

The Cramér-Rao lower bound approach allows one to include the spectrometer itself in consideration of the line fitting problem, and thus an exploration of the applicability of Cramér-Rao lower bounds to line fitting is also an exploration into the effect the spectrometer itself has on the information we can derive. An example of this can be found in Winick (1986) in which a Cramér-Rao lower bound argument is used to find the optimum pixel size such that the precision with which the location of a Gaussian emission line is maximized. Larger pixel sizes allow more photons to be collected and so the signal to noise ratio in each pixel is smaller. However, smaller pixels allow the line position to be better determined because there are more pixels per unit wavelength. The tradeoff between these two competing demands results in an optimum pixel size. Related questions of interest can be answered using the combination of an emission model, a spectrometer model and Cramér-Rao formalism.

2. THEORY

2.1. Maximum likelihood and the Cramér-Rao lower bound

A spectrometer observes an intensity emission profile \( D = \{D_1, D_2, \ldots, D_N\} \) over \( N \) channels at \( x = \{x_1, x_2, \ldots, x_N\} \). The features in the spectra,
such as emission lines, background emission and artifacts - are described using a total of \( N_v \) variables, where \( N_v < N_d \). Hence fitting the observed spectrum with \( N_v < N_d \) variables becomes an optimization problem. The variables form a set \( V \), and describe the number of counts per channel through an appropriate choice of function \( F(x, V) = F_i(V) \). It is up to the observer to pick an \( F \) and \( V \) that adequately describes the emission spectra for their particular purpose. The choice is complicated by the fact that many solar plasmas are inhomogenous in density, temperature, ionic composition, etc. In addition, the plasma is temporally inhomogeneous over the integration time of the exposure, and spatially inhomogeneous along the line of sight and across the field of view. The values found for the variables \( V \), are almost inevitably weighted averages of the true range of values in the observed emitting volume. This paper does not address the relation of the spectrum to the true physical state of the emitting volume. However, once the observer has chosen \( F(x, V) \) the precision with which the variables \( V \) can be determined is, as will be shown below, calculable. Discussion of the particular choice of \( F, V \) is deferred until Section 3.

The variable set \( V \) describes the spectrum observed. Multiple observations show that the data \( D_i \) in each channel is noisy, and since the spectrum is described by \( V \), the noise in each channel \( i \) can be described by some normalized probability distribution \( f_i(D_i|V) \). Assuming that the distribution function in each channel is independent of every other distribution function, the overall probability distribution function is therefore

\[
L(D, V) = f_1(D_1|V)f_2(D_2|V)\ldots f_{N_d}(D_{N_d}|V)
\]

The function \( L \) is known as the Likelihood Function (LF). The maximum likelihood principle directs us to choose the value of \( V \), \( V_0 \), say such that for any admissible value of \( V \),

\[
L(D, V_0) \geq L(D, \tilde{V})
\]

In practice it is common to use the log likelihood function to look for \( V = (V_1, \ldots, V_{N_v}) \) such that

\[
\frac{\partial \log L}{\partial V_j} = 0, 1 \leq j \leq N_v.
\]

\( \tilde{V} \) is a maximum of equation 1 if the matrix

\[
\left[ \frac{\partial^2 \log L}{\partial V_p \partial V_q} \right]_{\tilde{V}}
\]

is negative definite.

Take \( t \) as an unbiased estimator of some function of \( V \) say \( \tau(V) \). It can be shown that (Kendall & Stuart, 1973)

\[
\text{var}(t) \geq \sum_{p=1}^{N_v} \sum_{q=1}^{N_v} \frac{\partial \tau}{\partial V_p} \frac{\partial \tau}{\partial V_q} \mathbf{I}^{-1}
\]

where \( \mathbf{I} \) is the (Fisher) information matrix,

\[
I_{pq} = E \left( \frac{\partial \log L}{\partial V_p} \frac{\partial \log L}{\partial V_q} \right) = -E \left( \frac{\partial^2 \log L}{\partial V_p \partial V_q} \right)
\]

where \( E(.) \) denotes the expectation value of the parenthetical quantity (evaluation at the global maximum \( V_0 \)). Equation 5 is known as the Cramér-Rao inequality for a multivariate estimator. It expresses the fact that the variance \( \text{var}(t) \) of an unbiased estimator of the function \( \tau(V) \) is bounded below. If the function is set to one of the emission model variables then the theorem says that any unbiased estimator used to estimate the value of the variable has a minimum attainable variance. This implies that there is a limit to the precision of an experiment measuring the variables \( V \) via the emission model \( F(x, V) \). To evaluate equation 5, note that

\[
\left[ \frac{\partial^2 \log L}{\partial V_p \partial V_q} \right]_{V_0} = E \left( \frac{\partial \log L}{\partial V_p} \frac{\partial \log L}{\partial V_q} \right)
\]

that is, the expected value of \( \frac{\partial^2 \log L}{\partial V_p \partial V_q} \) occurs at the global maximum of the likelihood function. This defines a formally simple procedure for finding precision limits to an experiment. One finds \( V_0 \) and then evaluates equation 5 using equation 7. Finding \( V_0 \) is a global optimization problem and is, in general, nontrivial, although good estimates may be found by common line fitting routines. For demonstration purposes (Section 3), however, test cases are used in which the value of \( V_0 \) is known, i.e., everything is known about the emission. The remaining pieces of the framework needed to describe spectrometer experiments in a Cramér-Rao formalism are described below.

2.2. Emission model and Poisson noise

Attention is restricted to noise arising from the statistical nature of the emission process from the emitting region. Other sources of measurement error (both statistical and systematic) are not considered as one source of error is sufficient to demonstrate the applicability of the Cramér-Rao formalism. The emission in solar plasmas is assumed to follow Poisson statistics, and so if channel \( i \) measures \( D_i \) counts, then the distribution function \( f_i \) in each channel is

\[
f_i(D_i|V) = \frac{[F_i(V)]^{D_i}}{D_i!} \exp (-F_i(V))
\]

for some emission model \( F_i(V) \). Subsequently,

\[
\log L = \sum_{i=1}^{N_d} D_i \log [F_i(V)] - [F_i(V)] - \log (D_i!)
\]
and so the $p,q$th ($1 \leq p,q \leq N_v$) element of the information matrix is

$$I_{pq} = \sum_{i=1}^{N_d} \frac{1}{F_i(V)} \frac{\partial F_i(V)}{\partial V_p} \frac{\partial F_i(V)}{\partial V_q}$$

(10)

which is evaluated at $V = V_0$. The observer is interested in the value of emission model variables, $V_k$ ($1 \leq k \leq N_v$), and so $\tau = V_k$. Then any unbiased estimator $t$, of $V_k$, must have

$$\text{var}(t) \geq \left[ I^{-1} \right]_{kk}$$

(11)

where the right hand side of this equation is the Cramér-Rao lower bound to var($t$).

The emission model describes features in the spectrum. In solar physics, it is common to assume that the emission lines have a Gaussian shape and that the background emission has a quadratic dependence on wavelength (or its proxy, position on the detector $x = \{x_1, x_2, \ldots, x_{N_d}\}$). Hence,

$$F_i(V) = \mu + \nu x_i + \xi x_i^2 + \sum_{j=1}^{N_g} A_j \exp \left[ \frac{-(x_i - c_j)^2}{2\sigma_j^2} \right]$$

(12)

where $V = (\mu, \nu, \xi, A_1, c_1, \sigma_1, \ldots, A_{N_g}, c_{N_g}, \sigma_{N_g})$, the set of variables with which we describe the emission features in the spectrum. Values to the variables $V$ are found by fitting the model equation 12 to the data $D$, say by method of least squares, for example.

The final component in the description of an experimental setup is to model how the spectrometer is operated. To do this, the operation of the Coronal Diagnostic Spectrometer (CDS; Harrison et al. 1995) on board the Solar and Heliospheric Observatory is taken as a template of spectrometer operation. CDS is an extreme ultraviolet spectrometer covering wavelengths in the range 150 - 800 Å, using two spectrometer systems. The Normal Incidence Spectrometer (NIS) has two gratings that disperse two wavelength ranges, 308 - 381 Å (NIS1) and 513 - 633 Å (NIS2) (the Grazing Incidence Spectrometer observes spectra in 4 wavelength ranges 151 - 221 Å, 256 - 338 Å, 393 - 493 Å and 656 - 785 Å). Light is admitted to NIS via a slit, and images are built up by rastering the slit across the target. In each NIS study the entire wavelength range mentioned above is exposed to the CCD but only portions of the exposed CCD are stored due to memory requirements. This fixes a range to $x$ and data $D$ which is modeled by equation 12. The example application of the Cramér-Rao formalism is based on the Fe XVI 360.76Å emission observed by CDS in the execution of the SYNOP/v1 study. In this study, a window of emission from 360.123Å-361.386Å is observed using 18 pixels, and is modeled using a constant background emission ($\mu = 114$ photons/px) and a single Gaussian emission line (known as Gaussian 1; $A_1 = 1434$ photons/px, $\sigma_1 = 1.62$px) centered in the window ($c_1 = 0$px).

The applications below are not intended to model the precision behavior of CDS, but are intended to demonstrate the range of applicability of a Cramér-Rao approach to describing the precision aspects of general spectrometer experiments.

3. AN EXAMPLE APPLICATION

What is the effect on the achievable precision in measuring a Gaussian emission line when a smaller contaminating line appears in the wings of the larger line of interest? This can be modeled by considering Equation 12 when $N_g = 2$. The first Gaussian is that described in Section 2.2. The second Gaussian (the contaminating line, referred to as Gaussian 2) has amplitude 444 photons/px, a width of 0.5px and a position that varies across the main line ($\pm 5\sigma_1$). The background is assumed to be linear only. One approach to the problem of the contaminating line is to fit all 7 variables and so isolate the background and 2 Gaussians. Figure 1 shows the Cramér-Rao lower bounds (bold lines) to the achievable variance in the four parameters Gaussian amplitude $A_1$, $c_1$, $\sigma_1$ and $\mu$ as a function $c_2$, the location of the 2nd contaminating line. Normalized variances are also calculated via simulation using the CURVEFIT routine in IDL to fit a 1024 Poisson-noisy emission spectra (generated from Equation 12) at each value of $c_2$.

The contaminating Gaussian does of course cause the achievable variances to increase compared to the no contamination case. Note that the Cramér-Rao lower bounds vary with positions, a reflection of the interdependence of each Cramér-Rao lower bound on information available for each parameter, as expressed through Equation 6. The presence of the contaminating Gaussian can make a significant difference to the achievable precision. As expected, the contaminating Gaussian makes the smallest change to the precision when it is most distant from Gaussian (1). As the contaminating Gaussian moves towards the line centre, the achievable precisions become considerably higher than in the no contamination case.

Winkling out the contaminating line from the line of interest is a difficult problem, as evinced by the non-bold lines in Figure 1. The IDL routine CURVEFIT was used to generate the simulation-derived variances in fit parameter: the routine has a good performance when Gaussian (2) is far from the line centre, as one would expect. However, performance gets dramatically worse than the Cramér-Rao lower bound limit towards line centre. This means that there must be a large proportion of fit results that are relatively distant from the true value. Figures
Figure 1. Cramér-Rao lower bounds (bold lines) for the parameters of emission line (1); Gaussian amplitude $A_1$ (solid line) position $c_1$ (dashed), line width $\sigma_1$ (dot-dashed) and background $\mu$ (dot-dot-dot-dashed line). Also shown are normalised variances derived from simulation (non-bold lines). All variances are normalised to the Cramér-Rao lower bounds derived from having a single Gaussian (1) with no contaminating line.

2 and 3 show the distribution of simulation fit results as a function of the 2nd Gaussian position $c_2$. Only the background and $\chi^2$ do not exhibit a significant dependence with position. Although there are always some poorer fits regardless of the position of the 2nd Gaussian, CURVEFIT performs rather uniformly with regard to the $\chi^2$ measure. The routine appears to swap the identities of each line on occasion; sometimes the amplitude to Gaussian 1 is very small implying that Gaussian 2 must have a very large amplitude. Indeed, both amplitude populations show values more on the order of the other Gaussian. This leads to a possible mechanism which could cause the poor performance relative to the Cramér-Rao lower bounds calculated. The identity of each Gaussian, implicit in the Cramér-Rao lower bound calculation, is not present in CURVEFIT. The routine is designed to reduce the $\chi^2$ measure and so if changing what the user would see as the assignation of the line of interest with the contaminating line is not disallowed, then it is possible that it will do so. This could happen due a combination of noise and the local relative flatness of the search space. This can lead to fit values that appear to occur much more frequently than those expected by pure statistical noise, increasing the simulation-derived variances to those above the Cramér-Rao lower bound. Hence the fitting routine performs worse than Cramér-Rao lower bound limits due to the non-unique assignation of which fit parameters are associated with which emission line.

4. CONCLUSION

The aim of this paper is to show that a Cramér-Rao framework is a convenient way to pose questions about the precision of spectrometer experiments and line fitting. The example application chosen is relatively simple and is intended to be illustrative of the power and flexibility of this approach. The application of Cramér-Rao lower bound to models of solar emission spectra shows that it is possible to characterize the precision capabilities of spectrometers. The formalism also allows the exploration of the behaviour of fitting routines. One simple extension to this work would be to compare the Cramér-Rao lower bounds calculated here to those in a simulation where the 2nd smaller nuisance Gaussian is treated as a locally high level of noise. Such a comparison would allow one to judge whether it is better to fit the 2nd Gaussian or not from the point of view of the achievable precision on the parameters of the larger Gaussian of interest. Other applications will be described in Ireland (2004).

ACKNOWLEDGMENTS

This work is supported by NASA, contract number NNH04CC31C. The author wishes to thank N. H. Augustin, R. Fewster, A. D. Gordon for their helpful discussions and suggestions. This paper is dedicated to the memory of Andrew Lidsey.

REFERENCES


Rao, C.R., 1945, Bulletin of the Calcutta Mathematical Society, 37, 81


Figure 2. Fit values for the main Gaussian (1) and background as a function of the 2nd Gaussian position. The results are expressed as a difference from the true value in units of the average variance of the parameter, averaged over the range of 2nd Gaussian positions. The size of the average variance (in the relevant units) is shown in each plot title. The mean value at each value of 2nd Gaussian position is shown by the solid, non-horizontal line.
Figure 3. Fit values for the contaminating Gaussian (2) and chi-squared value as a function of the 2nd Gaussian position. The results are expressed as a difference from the true value in units of the average variance of the parameter, averaged over the range of 2nd Gaussian positions (the chi-squared values are shown as differences from 1). The size of the average variance (in the relevant units) is shown in each plot title. The mean parameter value at each value of 2nd Gaussian position is shown by the solid, non-horizontal line.