Comparison of Expected and Observed Fisher Information in Variance Calculations for Parameter Estimates

X. Cao and *J. C. Spall**†

*JHU Department of Applied Mathematics and Statistics, Baltimore, MD; and †JHU Applied Physics Laboratory, Laurel, MD

PL is responsible for building mathematical models for many defense systems and other types of systems. These models often require statistical estimation of parameters θ from data collected on the system. Maximum likelihood estimates (MLEs) are the most common type of statistical parameter estimates. Variance calculations and confidence intervals for MLEs are commonly used in system identification and statistical inference for the purpose of characterizing the uncertainty in the MLE.

Let $\hat{\theta}_n$ be an MLE of θ from a sample size of n.

To accurately construct such confidence intervals, one typically needs to know the variance of the MLE, $\text{var}(\hat{\theta}_n)$.

APPROACH

Standard statistical theory shows that the standardized MLE is asymptotically normally distributed with a mean of zero and the variance equal to a function of the Fisher information matrix (FIM) at the unknown parameter (see Section 13.3 in Ref. 1).

Two common approximations for the variance of the MLE are the inverse observed FIM (the same as the Hessian of the negative log-likelihood) and the inverse expected FIM,[2] both of which are evaluated at the MLE given sample data:

$$\bar{F}^{-1}(\hat{\theta}_n) \text{ or } \bar{H}^{-1}(\hat{\theta}_n),$$

where $\bar{F}(\hat{\theta}_n)$ is the average FIM at the MLE ("expected" FIM) and $\bar{H}(\hat{\theta}_n)$ is the average Hessian matrix at the MLE ("observed" FIM). The question we wish to answer is: *which of the two approximations above, the inverse expected FIM or the inverse observed FIM, is better for characterizing the uncertainty in the MLE?* In particular, the answer to this question applies to constructing confidence intervals for the MLE.

To answer the question, we find the variance estimate T that minimizes the mean-squared error (MSE):

$$\min_{T} E[(n\,\text{var}(\hat{\theta}_n) - T)^2].$$

In this work, T is constrained to be

$$\bar{F}^{-1}(\hat{\theta}_n) \text{ or } \bar{H}^{-1}(\hat{\theta}_n).$$

Under certain assumptions, if we ignore a term of magnitude $o(n^{-1})$, $\bar{F}^{-1}(\hat{\theta}_n)$ tends to outperform $\bar{H}^{-1}(\hat{\theta}_n)$ in estimating variance of normalized MLE.

That is:

$$E[(n\,\text{var}(\hat{\theta}_n) - \bar{F}^{-1}(\hat{\theta}_n))^2] < E[(n\,\text{var}(\hat{\theta}_n) - \bar{H}^{-1}(\hat{\theta}_n))^2].$$

Table 1. MSE criterion for observed and expected FIM.

Example	$\bar{F}^{-1}(\hat{\theta}_n)$ (Expected FIM)	$\bar{H}^{-1}(\hat{\theta}_n)$ (Observed FIM)
1	0.02077	0.02691
2	0.007636	0.008238

- We have conducted numerical examples on the signal-plus-noise problem.

- The estimation problem is the MLE for the variance of signal.

- Table 1 shows the values of the MSE criterion for two distinct examples (observed and expected Fisher information represents candidate variances of the MLE variance estimate).

CONCLUSION

The conclusion drawn from this work is that the expected Fisher information is better than the observed Fisher information (i.e., it has a lower MSE), as predicted by theory.

The bottom line of this work is that, under reasonable conditions, a variance approximation based on the expected FIM tends to outperform a variance approximation based on the observed FIM under an MSE criterion. This result suggests that, under certain conditions, the inverse expected FIM is a better estimate for the variance of MLE when used in confidence interval calculations.

These results, however, are preliminary. Significant additional work will be required to better relate the mathematical regularity conditions to practical problems of interest at APL and elsewhere and to extend the results to a multivariate parameter vector θ. Furthermore, to the extent that the preferred approximation remains the inverse expected FIM, it is expected that the Monte Carlo method of Spall[3] can be used to compute the expected FIM in complex problems of interest at APL and elsewhere. This work is ongoing as part of the first author's Ph.D. work at JHU.

[1]Spall, J. C., *Introduction to Stochastic Search and Optimization: Estimation, Simulation, and Control,* Wiley, Hoboken, NJ (2003).

[2]Efron, B., and Hinkley, D. V., "Assessing the Accuracy of the Maximum Likelihood Estimator: Observed Versus Expected Fisher Information," *Biometrika* 65(3), 457–482 (1978).

[3]Spall, J. C., "Monte Carlo Computation of the Fisher Information Matrix in Nonstandard Settings," *J. Comput. Graph. Stat.* 14(4), 889–909 (2005).