PROBING THE COVARIANCE MATRIX

Kenneth M. Hanson
Los Alamos National Laboratory, Los Alamos, NM 87545, USA
(e-mail: kmh@lanl.gov, http://public.lanl.gov/kmh/)

Abstract

Relationships between statistics and physics often provide a deeper understanding that can lead new or improved algorithmic approaches to solving statistics problems. It is well known that the negative logarithm of a probability distribution is analogous to a physical potential. Thus, $\varphi(\boldsymbol{a}) = -\log(p(\boldsymbol{a} \,|\, \boldsymbol{y}))$ is analogous to a potential, where $p(\boldsymbol{a} \,|\, \boldsymbol{y})$ is the posterior, vector $\boldsymbol{a}$ represents the n continuous parameters, and $\boldsymbol{y}$ represents the m measurements. The maximum a posteriori (MAP) solution, $\hat{\boldsymbol{a}}$, which minimizes $\varphi(\boldsymbol{a})$, is frequently chosen as the parameter estimator because it is easier to find than the posterior mean. In many inference problems the posterior can not be stated in analytic form, only evaluated by means of a computational model.

The inference process requires estimates of the uncertainties in , which are related to the width of the posterior, typically characterized in terms of the covariance matrix $\boldsymbol{C}$. Standard approaches to determining $\boldsymbol{C}$ include: 1) sensitivity analysis, 2) Markov chain Monte Carlo, and3) functional analysis. Each of these approaches has its advantages and disadvantages depending on the nature of the problem, for example, the magnitude of n and m and the cost of evaluating the forward model and its sensitivities.

I describe a novel alternative approach that may be advantageous in some situations. In the physics analogy, the notion is to determine the displacement of the equilibrium of the system ($\hat{\boldsymbol{a}}$) under the influence of an external force. The displacement is determined by the curvature (or stiffness) matrix describing the potential around $\hat{\boldsymbol{a}}$. In the inference problem, the idea is to add to $\varphi(\boldsymbol{a})$ a potential that is linear in $\boldsymbol{a}$ and find the new minimizer $\boldsymbol{a}'$. It is easy to show that $\Delta\boldsymbol{a} = \boldsymbol{a}' - \hat{\boldsymbol{a}} = \boldsymbol{C}\boldsymbol{f}$, where $\boldsymbol{f}$ is the force applied to the system; thus, the additional potential is $\boldsymbol{a}^T\boldsymbol{f}$. The force $\boldsymbol{f}$ represents a linear combination of the parameters about which we want to estimate the uncertainty. The variance in the direction of $\boldsymbol{f}$ is proportional to $\boldsymbol{f}^T\boldsymbol{C}\boldsymbol{f} = \boldsymbol{f}^T\Delta\boldsymbol{a}$. Furthermore, the covariance between $\boldsymbol{f}$ and another linear combination of parameters $\boldsymbol{g}$ is $\boldsymbol{g}^T\boldsymbol{C}\boldsymbol{f} = \boldsymbol{g}^T\Delta\boldsymbol{a}$. This approach to uncertainty estimation is most useful in situations in which a) the standard techniques are costly, b) it is relatively easy to find the minimum in $\varphi(\boldsymbol{a})$ and $\varphi(\boldsymbol{a}) + \boldsymbol{a}^T\boldsymbol{f}$, and c) one is interested in the uncertainty wrt. one or a few directions in the parameter space. The useful of this new technique is demonstrated with examples ranging from simple to complicated.

Key Words: covariance estimation, probability potential, posterior stiffness, linear response theory, dissipation-fluctuation relation