Title: UNCERTAINTY QUANTIFICATION FOR HOMELAND SECURITY APPLICATIONS

Author(s): Jeffrey A. Favorite

Submitted to: Talk for the LANL Uncertainty Quantification Working Group, Dec. 11, 2003
Uncertainty Quantification for Homeland Security Applications (U)

LA–UR–03–8914

Jeffrey A. Favorite
Los Alamos National Laboratory
X-5
Los Alamos, NM 87545

Uncertainty Quantification Working Group Seminar
December 11, 2003
Given a set of simulated gamma-ray leakage measurements and the unknown object model that best fit these measurements, standard tools of uncertainty quantification in data-fitting were applied to develop an estimate of the uncertainty of the dimensions in the model based on the statistical uncertainty of the measurements. At least two issues were discovered that need further exploration. One is that in standard data analysis, there are usually many more data points than there are model parameters to be fit, so that calculating the inverse of the Hessian (or curvature) matrix to obtain the covariance matrix is straightforward; in Homeland Security applications, however, there are more unknowns than data points, and it is not obvious how the correct covariance matrix should be obtained if the Hessian matrix is singular. Thus, in most problems of interest, it is still unclear how to relate the uncertainty in measurements to the uncertainty in model dimensions. A related issue is how the uncertainty in model dimensions should be used to obtain the uncertainty in other model quantities, such as material masses. The standard procedure is to use the variances and covariances of the dimensions in the simple error propagation formula; an alternative is to randomly sample the model dimensions from Gaussian distributions whose widths are related to the uncertainty in the dimensions, compute the associated masses, and then fit the mass distributions to Gaussians whose widths are related to the uncertainty in the masses. Both of these procedures require knowledge of the covariance matrix for the model dimensions.

Thus, a program to extend the standard methods of uncertainty quantification to develop an estimate of the covariance matrix in the case of a singular Hessian matrix would be of interest for Homeland Security applications.
Motivation

- Consider a border portal monitor (radiation detector) that “sees” a radioactive object. We want to use the data from the monitor to tell us how much radioactivity there is, with uncertainties.

We will start with an easier problem:

- Consider a radioactive object emitting $\gamma$ rays of discrete energies that are well resolved using high-purity germanium (HPGe) detectors
- We want to use $\gamma$ leakage measurements to tell us what the system is

![Diagram of a source, shield, and detector system with labeled interfaces]

- Notation:

  \[ M^g_o = \text{measured leakage for } \gamma \text{ line } g \ (g = 1, \ldots, G) \]
  \[ \sigma^g_o = \text{statistical uncertainty of } M^g_o \]
  \[ M^g = \text{calculated leakage for } \gamma \text{ line } g \ (g = 1, \ldots, G) \text{ for some } \textit{model} \]
  \[ \textit{model} = \text{a description of the system (materials, masses, interface locations, etc.), NOT a description of the } \gamma \text{ transport process} \]

- We assume that there is a single number (the $\gamma$ leakage) associated with a single energy line
Notation

- We consider only the transport of photons of discrete energies and assume that any scattered photons lose energy and are removed. The angular flux of photons at the discrete energy denoted by index $g$ is given by
  \[ \hat{\Omega} \cdot \vec{\nabla} \psi^g(r, \hat{\Omega}) + \Sigma^g_i(r) \psi^g(r, \hat{\Omega}) = q^g(r) \]
  for $g = 1, \ldots, G$. (This equation represents the forward problem.)

- The adjoint equation is
  \[ -\hat{\Omega} \cdot \vec{\nabla} \psi^{\ast g}(r, \hat{\Omega}) + \Sigma^g_i(r) \psi^{\ast g}(r, \hat{\Omega}) = q^{\ast g}(r) \]
  where the source is actually the detector response function.

- These equations can be rendered in operator notation as
  \[ L^g \psi^g = q^g \]
  and
  \[ L^{\ast g} \psi^{\ast g} = q^{\ast g} \]

- Suppose the scalar flux for each energy line $g$ is measured at a detector. The quantity of interest is
  \[ M^g = \int dV \int d\hat{\Omega} \Sigma^g_d(r) \psi^g(r, \hat{\Omega}) \]
  where the detector response function $\Sigma^g_d(r)$ is zero outside the detector volume.

- Introducing the inner product notation $\langle \cdot \rangle$ to mean an integral over all phase space (volume and angle), the quantity of interest is
  \[ M^g = \langle \Sigma^g_d \psi^g \rangle \]
  A weight function or detector efficiency can be built into $\Sigma^g_d(r)$. 
Strategies for optimization


  + The Schwinger method is iterative (i.e. implicit) but based on algebraic manipulations of the transport equation (i.e. explicit)

  + The method updates the unknown interface locations using

    \[
    R \Delta r = \frac{M^g - M_o^g}{M_o^g},
    \]

    where \( \Delta r \) is an \( N \times 1 \) vector and \( R \) is a \( G \times N \) matrix.

- Search schemes: Variational perturbation theory (Favorite) and a geometry-based scheme due to Diane Vaughan and Kevin Buescher (X-8)


  + It can be shown that

    \[
    \frac{\partial M^g}{\partial r_n} = \Delta q^g_n \int d\hat{\Omega} \psi^g(r_n) - \Delta \Sigma^g_{i,n} \int d\hat{\Omega} \psi^g(r_n) \psi^g(r_n),
    \]

    where \( \Delta q^g_n \) and \( \Delta \Sigma^g_{i,n} \) are the source and cross section differences across interface \( r_n \)

  + If, for each line, \( \epsilon^g \equiv \frac{1}{2} \left( \frac{M^g - M_o^g}{\sigma_o^g} \right)^2 \), then

    \[
    \frac{\partial \epsilon^g}{\partial r_n} = \frac{(M^g - M_o^g)}{(\sigma_o^g)^2} \frac{\partial M^g}{\partial r_n}
    \]

  + No numerical differentiation is required
Chi-squared and the covariance matrix

• $\chi^2$:

$$\chi^2 = \sum_{g=1}^{G} \left( \frac{M^g - M^o_g}{\sigma^g_o} \right)^2$$

• $\chi^2$ gradient vector:

$$\frac{\partial \chi^2}{\partial r_n} = 2 \sum_{g=1}^{G} \frac{M^g - M^o_g}{(\sigma^g_o)^2} \partial M^g \partial r_n$$

• Hessian (curvature) matrix of $\chi^2$:

$$\frac{\partial^2 \chi^2}{\partial r_n \partial r_m} = 2 \sum_{g=1}^{G} \left[ \frac{M^g - M^o_g}{(\sigma^g_o)^2} \partial^2 M^g \partial r_n \partial r_m + \frac{1}{(\sigma^g_o)^2} \partial M^g \partial M^g \partial r_m \partial r_n \right]$$

$$\approx 2 \sum_{g=1}^{G} \frac{1}{(\sigma^g_o)^2} \frac{\partial M^g}{\partial r_m} \frac{\partial M^g}{\partial r_n}$$

• Define the $\alpha$ matrix to be $\frac{1}{2}$ the Hessian matrix of $\chi^2$ (Press et al. call this the curvature in Chap. 15):

$$\left[ \alpha_{nm} \right] = \sum_{g=1}^{G} \frac{1}{(\sigma^g_o)^2} \frac{\partial M^g}{\partial r_m} \frac{\partial M^g}{\partial r_n}$$

• The covariance matrix of uncertainties in the estimated values of the interface locations is the inverse of $\alpha$ (the curvature?):

$$C = \alpha^{-1}$$

$$= \begin{bmatrix}
\sigma^2_{\eta_1} & \cdots & \sigma^2_{\eta_1,r_N} \\
\vdots & \ddots & \vdots \\
\sigma^2_{r_N,\eta_1} & \cdots & \sigma^2_{r_N}\n
\end{bmatrix}$$

• What if the Hessian is singular?
Relating uncertainty in interface locations to uncertainty in mass (1)

- Consider, for convenience, a spherically symmetric system

![Diagram of spherically symmetric system](image)

- What’s the uncertainty in the mass of region $n$?

$$m_n = \frac{4\pi}{3} \rho_n (r_n^3 - r_{n-1}^3)$$

- Standard formula for propagation of errors:

$$\sigma_{m_n}^2 = \left( \frac{\partial m_n}{\partial r_n} \right)^2 \sigma_{r_n}^2 + \left( \frac{\partial m_n}{\partial r_{n-1}} \right)^2 \sigma_{r_{n-1}}^2 + 2 \frac{\partial m_n}{\partial r_n} \frac{\partial m_n}{\partial r_{n-1}} \sigma_{r_n} \sigma_{r_{n-1}}$$

$$= \left( 4\pi \rho_n r_n^2 \right)^2 \sigma_{r_n}^2 + \left( -4\pi \rho_n r_{n-1}^2 \right)^2 \sigma_{r_{n-1}}^2 + 2 \left( 4\pi \rho_n r_n^2 \right) \left( -4\pi \rho_n r_{n-1}^2 \right) \sigma_{r_n} \sigma_{r_{n-1}}$$

$$= \left( 4\pi \rho_n \right)^2 \left( r_n^4 \sigma_{r_n}^2 + r_{n-1}^4 \sigma_{r_{n-1}}^2 - 2r_n^2 r_{n-1}^2 \sigma_{r_n} \sigma_{r_{n-1}} \right)$$
Relating uncertainty in interface locations to uncertainty in mass (2)

- Use the covariance matrix to randomly sample interface locations \( r_n \) and \( r_{n-1} \) from Gaussian distributions with the proper widths

- Use these to compute masses

- The masses should be distributed in a Gaussian whose width is \( \sigma_m \)

- What is the proper distribution for the interface locations?

\[
\Delta r = C^{\frac{1}{2}} \xi ,
\]

where \( \xi \) represents \( N \) independent random numbers drawn from a Gaussian distribution with mean 0 and half-width 1

- What is the square root of the covariance matrix? Use singular value decomposition (SVD) on the Hessian:

\[
\alpha = U W V^T ,
\]

where \( W = \begin{bmatrix} w_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & w_N \end{bmatrix} \) is the diagonal matrix of singular values. Now

\[
C = \alpha^{-1} = V \left[ \text{diag} \left( \frac{1}{w_n} \right) \right] U^T
\]

and

\[
C^{\frac{1}{2}} = V \left[ \text{diag} \left( \frac{1}{\sqrt{w_n}} \right) \right] U^T
\]

- If \( \alpha \) is not singular then \( U = V \); if \( \alpha \) is singular then \( U \neq V \)
Relating uncertainty in interface locations to uncertainty in leakage

• Use the covariance matrix to randomly sample interface locations \( r_n \) and \( r_{n-1} \) from Gaussian distributions with the proper widths

• Use these to compute leakage, \( M^g \)

• The leakages should be distributed in a Gaussian whose width is \( \sigma_o^g \)

OR

• Use the standard formula for propagation of errors:

\[
(\sigma^g)^2 = \sum_n \left( \frac{\partial M^g}{\partial r_n} \right)^2 \sigma_{r_n}^2 + 2\sum_{m>n} \sum_n \frac{\partial M^g}{\partial r_n} \frac{\partial M^g}{\partial r_m} \sigma_{r_n r_m}^2
\]

• \( \sigma^g \) should equal \( \sigma_o^g \)
Test problem

- Godiva (HEU) model:
  - Spherical
  - There are four uranium $\gamma$ lines, but only three can escape this model

- $\gamma$ leakage from a Monte Carlo calculation:

<table>
<thead>
<tr>
<th>Energy (keV)</th>
<th>$\gamma$ leakage ($s^{-1}$) and 1σ uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>144</td>
<td>Not observed</td>
</tr>
<tr>
<td>186</td>
<td>$5.28 \times 10^3 \pm 40.82%$</td>
</tr>
<tr>
<td>766</td>
<td>$2.50 \times 10^3 \pm 0.41%$</td>
</tr>
<tr>
<td>1001</td>
<td>$1.01 \times 10^4 \pm 0.33%$</td>
</tr>
</tbody>
</table>

- A one-dimensional deterministic $S_N$ code, PARTISN, was used in the optimization process
  + $S_{32}$
  + No scattering
  + Discrete-energy total cross sections from the Monte Carlo library

- The optimization details are interesting but not important here; assume that I’ve found the minimum $\chi^2$ for each case
Test problem

Estimating $\sigma^g$ for the line leakages
Test problem 1: Two shield unknowns

- **Godiva model:**
- The assumed $r_1$ and $r_4$ were not correct; this made little difference

![Diagram of shields and distances]

- Actual: 8.741  12.4  12.9  13.2  

- **Standard deviation in line leakages (expressed as relative errors):**

<table>
<thead>
<tr>
<th>Line (keV)</th>
<th>$\sigma^g_\omega$</th>
<th>$\sigma^g$ from error prop.</th>
<th>$\sigma^g$ from Gauss. fit</th>
</tr>
</thead>
<tbody>
<tr>
<td>186</td>
<td>0.408</td>
<td>0.316</td>
<td>0.299</td>
</tr>
<tr>
<td>766</td>
<td>0.0041</td>
<td>0.0034</td>
<td>0.0032</td>
</tr>
<tr>
<td>1001</td>
<td>0.0033</td>
<td>0.0028</td>
<td>0.0027</td>
</tr>
</tbody>
</table>

$\sigma^g$ from error propagation and Gaussian fit are similar but smaller than $\sigma^g_\omega$.

- **186-keV line fit**
- **1001-keV line fit**
Test problem 2: One shield, one source unknown

- Godiva model:

  - Standard deviation in line leakages (expressed as relative errors):

<table>
<thead>
<tr>
<th>Line (keV)</th>
<th>$\sigma_{\text{g}}$</th>
<th>$\sigma_{\text{g}}$ from error prop.</th>
<th>$\sigma_{\text{g}}$ from Gauss. fit</th>
</tr>
</thead>
<tbody>
<tr>
<td>186</td>
<td>0.408</td>
<td>0.218</td>
<td>0.204</td>
</tr>
<tr>
<td>766</td>
<td>0.0041</td>
<td>0.0037</td>
<td>0.0033</td>
</tr>
<tr>
<td>1001</td>
<td>0.0033</td>
<td>0.0031</td>
<td>0.0033</td>
</tr>
</tbody>
</table>

$\sigma_{\text{g}}$ from error propagation and Gaussian fit are similar but smaller than $\sigma_{\text{o}}^{g}$.

- 186-keV line fit

- 1001-keV line fit
Test problem 3: All four radii unknown

- Godiva model:

- Note: The Hessian is singular for this problem.

- Nevertheless, …

- Standard deviation in line leakages (expressed as relative errors):

<table>
<thead>
<tr>
<th>Line (keV)</th>
<th>( \sigma^g )</th>
<th>( \sigma^g ) from error prop.</th>
<th>( \sigma^g ) from Gauss. fit</th>
</tr>
</thead>
<tbody>
<tr>
<td>186</td>
<td>0.408</td>
<td>0.408</td>
<td>0.296</td>
</tr>
<tr>
<td>766</td>
<td>0.0041</td>
<td>0.0041</td>
<td>0.0036</td>
</tr>
<tr>
<td>1001</td>
<td>0.0033</td>
<td>0.0033</td>
<td>0.0033</td>
</tr>
</tbody>
</table>

\( \sigma^g \) from error propagation is correct! \( \sigma^g \) from Gaussian fit is smaller than \( \sigma^g \).

- 186-keV line fit

- 1001-keV line fit
Test problem

Estimating $\sigma_m$ for material masses
Test problem 1: Two shield unknowns

- **Godiva model:**

  ![Godiva model diagram]

  - Actual: 8.741
  - Assumed: 8.7046
  - Void
  - Lead
  - Aluminum

<table>
<thead>
<tr>
<th>Shell</th>
<th>Mass (kg)</th>
<th>$\sigma_m$ (kg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lead</td>
<td>10.950</td>
<td>0.797</td>
</tr>
<tr>
<td>Al</td>
<td>1.941</td>
<td>0.693</td>
</tr>
</tbody>
</table>

- **Mass and standard deviation (from the error propagation formula):**

- **Mass and standard deviation (from Gaussian fit):**

<table>
<thead>
<tr>
<th>Shell</th>
<th>Mass (kg)</th>
<th>$\sigma_m$ (kg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lead</td>
<td>10.963</td>
<td>0.783</td>
</tr>
<tr>
<td>Al</td>
<td>1.931</td>
<td>0.683</td>
</tr>
</tbody>
</table>

- Error propagation formula and Gaussian fit yield similar results.

- **Lead shell fit**

  ![Lead shell fit graph]

  - 0 500 1000 1500 2000 2500 3000 3500 4000 4500 5000

- **Aluminum shell fit**

  ![Aluminum shell fit graph]

  - 0 500 1000 1500 2000 2500 3000 3500 4000 4500 5000
Test problem 2: One shield, one source unknown

- **Godiva model:**

- **Mass and standard deviation (from the error propagation formula):**

<table>
<thead>
<tr>
<th>Shell</th>
<th>Mass (kg)</th>
<th>$\sigma_m$ (kg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Godiva</td>
<td>51.303</td>
<td>1.132</td>
</tr>
<tr>
<td>Lead</td>
<td>11.358</td>
<td>0.388</td>
</tr>
<tr>
<td>Al</td>
<td>1.759</td>
<td>0.092</td>
</tr>
</tbody>
</table>

- **Mass and standard deviation (from Gaussian fit):**

<table>
<thead>
<tr>
<th>Shell</th>
<th>Mass (kg)</th>
<th>$\sigma_m$ (kg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Godiva</td>
<td>51.316</td>
<td>1.125</td>
</tr>
<tr>
<td>Lead</td>
<td>11.361</td>
<td>0.382</td>
</tr>
<tr>
<td>Al</td>
<td>1.758</td>
<td>0.091</td>
</tr>
</tbody>
</table>

- Error propagation formula and Gaussian fit yield similar results.

- **Godiva fit**

- **Lead shell fit**
Test problem 3: All four radii unknown (1)

- Godiva model:

- Note: The Hessian is singular for this problem.

Actual: 8.741  
Assumed: 8.7042

- Mass and standard deviation (from the error propagation formula):

<table>
<thead>
<tr>
<th>Shell</th>
<th>Mass (kg)</th>
<th>$\sigma_m$ (kg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Godiva</td>
<td>52.839</td>
<td>2.698</td>
</tr>
<tr>
<td>Lead</td>
<td>10.637</td>
<td>0.782</td>
</tr>
<tr>
<td>Al</td>
<td>1.966</td>
<td>2.007</td>
</tr>
</tbody>
</table>

- Mass and standard deviation (from Gaussian fit):

<table>
<thead>
<tr>
<th>Shell</th>
<th>Mass (kg)</th>
<th>$\sigma_m$ (kg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Godiva</td>
<td>51.765</td>
<td>3.543</td>
</tr>
<tr>
<td>Lead</td>
<td>10.951</td>
<td>1.043</td>
</tr>
<tr>
<td>Al</td>
<td>1.936</td>
<td>2.004</td>
</tr>
</tbody>
</table>

- Which $\sigma_m$ is correct? (Recall that err. prop. gave correct $\sigma^g$ for lines.)

- Godiva fit

- Aluminum shell fit
Test problem 3: All four radii unknown (2)

- **Godiva model:**

- Note: The Hessian is singular for this problem.

<table>
<thead>
<tr>
<th>Mass and standard deviation (from the error propagation formula):</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shell</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>Godiva</td>
</tr>
<tr>
<td>Lead</td>
</tr>
<tr>
<td>Al</td>
</tr>
</tbody>
</table>

- **Gaussian fit:**

<table>
<thead>
<tr>
<th>Shell</th>
<th>Mass (kg)</th>
<th>$\sigma_m$ (kg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Godiva</td>
<td>52.839</td>
<td>2.698</td>
</tr>
<tr>
<td>Lead</td>
<td>10.637</td>
<td>0.782</td>
</tr>
<tr>
<td>Al</td>
<td>1.966</td>
<td>2.007</td>
</tr>
</tbody>
</table>

- **Gaussian fit with mean fixed:**

<table>
<thead>
<tr>
<th>Shell</th>
<th>Mass (kg)</th>
<th>$\sigma_m$ (kg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Godiva</td>
<td>51.765</td>
<td>2.264</td>
</tr>
<tr>
<td>Lead</td>
<td>10.951</td>
<td>0.662</td>
</tr>
<tr>
<td>Al</td>
<td>1.936</td>
<td>2.022</td>
</tr>
</tbody>
</table>

- The fit should be two-sided.

- **Godiva fit**

- **Lead shell fit**
Summary and conclusions

• We have just begun to apply standard methods of uncertainty quantification to problems of concern to Homeland Security

• The work presented here was performed on a very small budget in order to understand the standard methods

• Numerical tests were run on a problem of interest, but this problem is much easier than the general portal monitor problem

• Areas in need of further research have been identified

  + For $\sigma_m$ (statistical uncertainty in estimated mass), which is correct: the error propagation formula or Gaussian sampling of the unknown radii?

  + Why is $\sigma_g$ (statistical uncertainty in line leakage estimated using covariance matrix of unknown radii) accurately calculated from the error propagation formula when the Hessian is singular, but not when the Hessian is not singular?

  + Why is $\sigma_g$ not accurately calculated from Gaussian sampling of the unknown radii?

  + What are the implications of using SVD to invert a singular Hessian?

• Can these methods be applied to the more general portal monitor problem? (What does it mean to minimize $\chi^2$ in such problems?)