Marginalizing unknowns in GULIPS

Mikko Orispää

June 7, 2005

Let us consider a (real) equation

$$Ax = m + \varepsilon, \tag{1}$$

where $A \in \mathbb{R}^{(m \times n)}$ is called *the theory matrix* , $m \in \mathbb{R}^m$ is *the measurement*, $\varepsilon \in \mathbb{R}^m$ is *the error* with *the covariance matrix* $\Sigma = \langle \varepsilon \cdot \varepsilon^T \rangle$ and $x \in \mathbb{R}^n$ is *the unknown.* GULIPS transforms the equation (1) in the form

$$Rx = Y + \varepsilon' \tag{2}$$

where $R \in \mathbb{R}^{(n \times n)}$ is so called *target matrix* and $Y \in \mathbb{R}^n$ is *target vector.* The covariance matrix of the error ε' is identity matrix. The target matrix R is upper triangular, so solving the equation (2) is numerically cheap. In GULIPS the posteriori covariance matrix can be obtained easily by

$$\Sigma_{ps} = (R^T R)^{-1} = R^{-1}(R^{-1})^T \tag{3}$$

Now let us write the unknown vector x in the form

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \tag{4}$$

where $x_1 \in \mathbb{R}^{n'}$ and $x_2 \in \mathbb{R}^{n''}$ and respectively the covariance matrix in the form

$$\Sigma_{ps} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}, \tag{5}$$

where $\Sigma_{11} \in \mathbb{R}^{n' \times n'}$, $\Sigma_{12} \in \mathbb{R}^{n' \times n''}$, $\Sigma_{21} \in \mathbb{R}^{n'' \times n'}$ and $\Sigma_{22} \in \mathbb{R}^{n'' \times n''}$. Our intention is to marginalize away the x_1 part of the unknown.

Now,

$$p(x_2) = Ce^{-\frac{1}{2}(x_2 - \mu)^T \Sigma_{22}^{-1} (x_2 - \mu)}, \tag{6}$$

where μ is the mean (or expected value). In other words, we only need the Σ_{22} part of the covariance matrix.

Since R is upper triangular, R^{-1} is also upper triangular. Let us write R and R^{-1} as

$$R = \begin{pmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{pmatrix} \text{ and } R^{-1} = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}. \tag{7}$$

Then from eq. (3) we get

$$\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix} \begin{pmatrix} A_{11}^T & 0 \\ A_{12}^T & A_{22}^T \end{pmatrix} \tag{8}$$

$$= \begin{pmatrix} A_{11}A_{11}^T + A_{12}A_{12}^T & A_{12}A_{22}^T \\ A_{22}A_{12}^T & A_{22}A_{22}^T \end{pmatrix}. \tag{9}$$

Hence,

$$\Sigma_{22} = A_{22}A_{22}^T. \tag{10}$$

It is also easy to show that

$$R_{22}A_{22} = A_{22}R_{22} = I, \tag{11}$$

and therefore

$$R_{22}^{-1} = A_{22}. \tag{12}$$

So finally we get that

$$\Sigma_{22} = A_{22}A_{22}^T = R_{22}^{-1}(R_{22}^{-1})^T = (R_{22}^T R_{22})^{-1}, \tag{13}$$

where R_{22} is upper triangular. Moreover, from eq. (2) we get

$$\begin{pmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}, \tag{14}$$

which implies

$$R_{22}x_2 = y_2 + \varepsilon_2. \tag{15}$$

To summarize: if we want to marginalize away x_1 from the equation

$$R \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix},$$

where R is upper triangular,

$$R = \begin{pmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{pmatrix},$$

and the posteriori covariance matrix Σ is written as

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

we have to only take R_{22} part of R, Σ_{22} part of Σ and y_2 part of Y, because then

$$R_{22}x_2 = y_2 + \varepsilon_2,$$

and the covariance matrix is $\Sigma_{22} = (R_{22}^T R_{22})^{-1}$.