Useful Formulas and Some Details

Rui Paulo, paraphrased by Robert Wolpert

Originally: January 30 2004

1 Likelihood Expressions

The design set for the field data is $D^F = \{x_1, \ldots, x_\ell\}$. At input vector x_j a total of n_j replicates are measured, denoted by $y_k^F(x_j)$ for $1 \le k \le n_j$. The sufficient statistics are

$$\bar{y}^F = (\bar{y}^F(x_j),\ j = 1, \ldots, \ell)'$$
$$s_F^2 = \sum_{j=1}^{\ell} \sum_{k=1}^{n_j} \left(y_k^F(x_j) - \bar{y}^F(x_j) \right)^2,$$

where $\bar{y}^F(x_j) := \sum_{k=1}^{n_j} y_k^F(x_j)/n_j$. Let

- $\boldsymbol{y}^M$ = code data, *i.e.*, computer code observed at design set D^M;
- $\boldsymbol{u}_\star$ = "true" value of the calibration parameter;
- $\boldsymbol{y}_\star^M$ = code observed at the set D^F augmented with the true value of the calibration parameter, $\boldsymbol{u}_\star$, which we will denote by $D_\star^F$;
- $\boldsymbol{b}$ = discrepancy (or "bias") observed at $D_\star^F$.

Consequence of the modeling strategy:

$$\bar{y}^F = \boldsymbol{y}_\star^M + \boldsymbol{b} + \bar{\epsilon}, \quad \epsilon \mid \lambda^F \sim \mathsf{No}(0, \boldsymbol{\Sigma}^F), \quad \boldsymbol{\Sigma}^F = \operatorname{diag} \boldsymbol{n}^{-1}/\lambda^F, \tag{1}$$

which in turn implies that, independently,[1]

$$\bar{y}^F \mid \boldsymbol{y}_\star^M, \boldsymbol{b}, \lambda^F \sim \mathsf{No}(\boldsymbol{y}_\star^M + \boldsymbol{b}, \boldsymbol{\Sigma}^F) \tag{2}$$
$$s_F^2 \mid \lambda^F \sim \frac{1}{\lambda^F} \chi^2 \left(\textstyle\sum_{i=1}^{\ell} (n_i - 1) \right). \tag{3}$$

[1] Note that $\sum_{i=1}^{\ell} (n_i - 1) = (n_+ - \ell)$

Also, it is clear that

$$
\begin{aligned}
\boldsymbol{b} \mid \theta^b, \mu^b \quad & \sim \mathsf{No}(\mathbf{1}\mu^b, \quad c^b(D^F, D^F)/\lambda^b) && \equiv \mathsf{No}(\boldsymbol{\mu}^b, \boldsymbol{\Sigma}^b) && (4)\\
\boldsymbol{y}^M \mid \theta^L, \theta^M \quad & \sim \mathsf{No}(\boldsymbol{X}\theta^L, \; c^M(D^M, D^M)/\lambda^M) && \equiv \mathsf{No}(\boldsymbol{\mu}^M, \boldsymbol{\Sigma}^M) && (5)\\
\boldsymbol{y}^M_\star \mid \theta^L, \theta^M, \boldsymbol{u}_\star \quad & \sim \mathsf{No}(\boldsymbol{X}_\star\theta^L, c^M(D^F_\star, D^F_\star)/\lambda^M) && \equiv \mathsf{No}(\boldsymbol{\mu}^M_\star, \boldsymbol{\Sigma}^M_\star) && (6)\\
\boldsymbol{y}^M_\star \mid \boldsymbol{y}^M, \theta^L, \theta^M, \boldsymbol{u}_\star & \sim && \mathsf{No}(\boldsymbol{\mu}_{\star|\bullet}, \boldsymbol{\Sigma}^M_{\star|\bullet}) && (7)
\end{aligned}
$$

where $\boldsymbol{\mu}_{\star|\bullet}$ and $\boldsymbol{\Sigma}^M_{\star|\bullet}$ ("field estimates conditional on design") are given by

$$\boldsymbol{\mu}_{\star|\bullet} = \boldsymbol{\mu}^M_\star + \boldsymbol{\Sigma}_{\star\bullet}[\boldsymbol{\Sigma}^M]^{-1}\big(\boldsymbol{y}^M - \boldsymbol{\mu}^M\big) \tag{8}$$

$$\boldsymbol{\Sigma}_{\star|\bullet} = \boldsymbol{\Sigma}^M_\star - \boldsymbol{\Sigma}_{\star\bullet}[\boldsymbol{\Sigma}^M]^{-1}\boldsymbol{\Sigma}_{\bullet\star} \tag{9}$$

with

$$\boldsymbol{\Sigma}_{\star\bullet} = c^M(D^F_\star, D^M)/\lambda^M \tag{10}$$

$$\boldsymbol{\Sigma}_{\bullet\star} = \boldsymbol{\Sigma}_{\star\bullet}{}' = c^M(D^M, D^F_\star)/\lambda^M \tag{11}$$

With the above in mind, we have

$$f(\bar{y}^F, s^2_F, \boldsymbol{b}, \boldsymbol{y}^M_\star, \boldsymbol{y}^M \mid \underbrace{\theta^L, \theta^M, \mu^b, \theta^b, \lambda^F, \boldsymbol{u}_\star}_{\theta}) = \tag{12}$$

$$
\begin{aligned}
=& f(\bar{y}^F, s^2_F \mid \boldsymbol{b}, \boldsymbol{y}^M_\star, \boldsymbol{y}^M, \theta)\times && (13)\\
& f(\boldsymbol{b} \mid \boldsymbol{y}^M_\star, \boldsymbol{y}^M, \theta)\times && (14)\\
& f(\boldsymbol{y}^M_\star \mid \boldsymbol{y}^M, \theta)\times && (15)\\
& f(\boldsymbol{y}^M \mid \theta) && (16)
\end{aligned}
$$

$$
\begin{aligned}
=& f(s^2_F \mid \lambda^F)\times && (17)\\
& f(\bar{y}^F \mid \boldsymbol{b}, \boldsymbol{y}^M_\star, \lambda^F)\times && (18)\\
& f(\boldsymbol{b} \mid \theta^b, \mu^b)\times && (19)\\
& f(\boldsymbol{y}^M_\star \mid \boldsymbol{y}^M, \theta^L, \theta^M, \boldsymbol{u}_\star)\times && (20)\\
& f(\boldsymbol{y}^M \mid \theta^L, \theta^M). && (21)
\end{aligned}
$$

Note how we know all these densities, and that the last four are all multivariate Gaussian. For that reason, we are actually able to integrate out $\boldsymbol{y}_\star^M$ and $\boldsymbol{b}$ in closed form to get

$$f(\bar{y}^F, s_F^2, \boldsymbol{y}^M \mid \theta) = \lambda^F \ \chi^2\big(\lambda^F \ s_F^2 \mid \textstyle\sum_{i=1}^{\ell}(n_i - 1)\big)\times \tag{22}$$
$$\mathsf{No}\big(\bar{y}^F \mid \boldsymbol{\mu}_{\star|\bullet} + b, \boldsymbol{\Sigma}_{\star|\bullet} + \boldsymbol{\Sigma}^F\big)\times \tag{23}$$
$$\mathsf{No}(\boldsymbol{y}^M \mid \boldsymbol{\mu}^M, \boldsymbol{\Sigma}^M). \tag{24}$$

We can also marginalize only over $\boldsymbol{y}_\star^M$ to get:

$$f(\bar{y}^F, s_F^2, \boldsymbol{y}^M, \boldsymbol{b} \mid \theta) = \lambda^F \ \chi^2\big(\lambda^F \ s_F^2 \mid \textstyle\sum_{i=1}^{\ell}(n_i - 1)\big)\times \tag{25}$$
$$\mathsf{No}\big(\bar{y}^F \mid \boldsymbol{\mu}_{\star|\bullet} + b, \boldsymbol{\Sigma}_{\star|\bullet} + \boldsymbol{\Sigma}^F\big)\times \tag{26}$$
$$\mathsf{No}\big(\boldsymbol{b} \mid \boldsymbol{\mu}^b, \boldsymbol{\Sigma}^b\big)\times \tag{27}$$
$$\mathsf{No}(\boldsymbol{y}^M \mid \boldsymbol{\mu}^M, \boldsymbol{\Sigma}^M), \tag{28}$$

or only over $\boldsymbol{b}$ for

$$f(\bar{y}^F, s_F^2, \boldsymbol{y}^M, \boldsymbol{y}_\star^M \mid \theta) = \lambda^F \ \chi^2\big(\lambda^F \ s_F^2 \mid \textstyle\sum_{i=1}^{\ell}(n_i - 1)\big)\times \tag{29}$$
$$\mathsf{No}\big(\bar{y}^F \mid \boldsymbol{y}_\star^M + \boldsymbol{\mu}^b, \boldsymbol{\Sigma}^b + \boldsymbol{\Sigma}^F\big)\times \tag{30}$$
$$\mathsf{No}\big(\boldsymbol{y}_\star^M \mid \boldsymbol{\mu}_{\star|\bullet}, \boldsymbol{\Sigma}_{\star|\bullet}\big)\times \tag{31}$$
$$\mathsf{No}(\boldsymbol{y}^M \mid \boldsymbol{\mu}^M, \boldsymbol{\Sigma}^M). \tag{32}$$

In the second stage of the modular approach, the factor $\mathsf{No}(\boldsymbol{y}^M \mid \boldsymbol{\mu}^M, \boldsymbol{\Sigma}^M)$ of (32) is omitted. Another alternative to these expressions is the one that arises from considering the joint distribution directly:

$$f(\bar{y}^F, s_F^2, \boldsymbol{y}^M \mid \theta) = \lambda^F \chi^2\left(\lambda^F \ s_F^2 \,\middle|\, \sum_{i=1}^{\ell}(n_i - 1)\right)\times \tag{33}$$
$$\mathsf{No}\big((\boldsymbol{y}^{M\prime}, \boldsymbol{y}^{F\prime})' \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}\big), \tag{34}$$

where

$$\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{X} & \boldsymbol{0} \\ \boldsymbol{X}_\star & \boldsymbol{1} \end{bmatrix} \begin{pmatrix} \theta^L \\ \mu^b \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}^M & \boldsymbol{\Sigma}_{\bullet\star} \\ \boldsymbol{\Sigma}_{\star\bullet} & \boldsymbol{\Sigma}_\star^M + \boldsymbol{\Sigma}^b + \boldsymbol{\Sigma}^F \end{bmatrix}. \tag{35}$$

2 Second stage MCMC if all parameters, except precisions, are fixed

This is what **?** implemented in the code sent to GM in 2004, with $\mu^b \equiv 0$:

1. $[\boldsymbol{y}_\star^M, \boldsymbol{b} \mid \lambda^b, \lambda^F, \boldsymbol{u}_\star, \boldsymbol{y}^M, \bar{y}^F, s_F^2] \sim$ Kalman Filter, see details in Section (3) below.

2. $[\lambda^F \mid \lambda^b, \boldsymbol{u}_\star, \boldsymbol{y}^M, \boldsymbol{b}, \boldsymbol{y}^M, \bar{y}^F, s_F^2] \sim \mathsf{Ga}(\alpha_{F|\bullet}, r_{F|\bullet})$, gamma with conditional shape and rate parameters

$$\begin{aligned}\alpha_{F|\bullet} &= \alpha_F + \sum n_i/2 \\ r_{F|\bullet} &= r_F + s_F^2/2 + (\bar{y}^F - \boldsymbol{b} - \boldsymbol{y}_\star^M)' \; [\operatorname{diag} \boldsymbol{n}] \; (\bar{y}^F - \boldsymbol{b} - \boldsymbol{y}_\star^M)/2\end{aligned}$$

   and, *a priori*, $\lambda^F \sim \mathsf{Ga}(\alpha_F, r_F)$.

3. $[\lambda^b, \boldsymbol{u}_\star \mid \lambda^F, \boldsymbol{y}_\star^M, \boldsymbol{b}, \boldsymbol{y}^M, \bar{y}^F, s_F^2] \propto \pi(\lambda^b)\, f(\boldsymbol{b} \mid \theta^b)\, f(\boldsymbol{y}_\star^M \mid \boldsymbol{y}^M, \theta^L, \theta^M, \boldsymbol{u}_\star)$.

   This vector is sampled jointly using a Metropolis step. The proposal is the full conditional of λ^b times the prior on $\boldsymbol{u}_\star$.

   The full conditional is $[\lambda^b \mid \lambda^F, \boldsymbol{u}_\star, \boldsymbol{y}_\star^M, \boldsymbol{b}, \bar{y}^F, \boldsymbol{y}^M, s_F^2] \sim \mathsf{Ga}(\alpha_{b|\bullet}, r_{b|\bullet})$, gamma with shape and rate parameters

$$\begin{aligned}\alpha_{b|\bullet} &= \alpha_b + \ell/2 \\ r_{b|\bullet} &= r_b + \boldsymbol{b}'[c^b(D^F, D^F)]^{-1}\boldsymbol{b}/2\end{aligned}$$

   and, *a priori*, $\lambda^b \sim \mathsf{Ga}(\alpha_b, r_b)$.

   This step can also be done by sampling λ^b directly from its full conditional, followed by a Metropolis step to sample $\boldsymbol{u}_\star$ from its full conditional. A proposal can be the prior itself, if that works. In the example described in the software documentation, it did not make a noticeable difference whether we were doing this or sampling from the joint.

3 Note on the Kalman Filter part

All statements are conditional on the parameters. By sufficiency,

$$f(\boldsymbol{y}_\star^M, \boldsymbol{b} \mid \boldsymbol{y}^M, y^F) = f(\boldsymbol{y}_\star^M, \boldsymbol{b} \mid \boldsymbol{y}^M, \bar{y}^F)$$

Also, it is clear that

$$\boldsymbol{y}^M, \bar{y}^F, \boldsymbol{y}_\star^M, \boldsymbol{b} \sim \mathsf{No}(\boldsymbol{\mu}, \boldsymbol{\Sigma}),$$

where

$$\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{X} & \boldsymbol{0} \\ \boldsymbol{X}_\star & \boldsymbol{1} \\ \boldsymbol{X}_\star & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{1} \end{bmatrix} \begin{pmatrix} \theta^L \\ \mu^b \end{pmatrix}$$

and

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}^M & & & \\ \boldsymbol{\Sigma}_{\star\bullet} & \boldsymbol{\Sigma}_\star^M + \boldsymbol{\Sigma}^b + \boldsymbol{\Sigma}^F & & \\ \boldsymbol{\Sigma}_{\star\bullet} & \boldsymbol{\Sigma}_\star^M & \boldsymbol{\Sigma}_\star^M & \\ \boldsymbol{0} & \boldsymbol{\Sigma}^b & \boldsymbol{0} & \boldsymbol{\Sigma}^b \end{bmatrix}$$

so that it's easy to compute the mean and covariance of the conditional distribution. How can one partition the conditional density of $(\boldsymbol{y}_\star^M, \boldsymbol{b})$ given the data? The conditionals and marginals are all Gaussian with mean and covariance that follow the pattern

$$\begin{aligned} \boldsymbol{\Sigma} &= (A^{-1} + B^{-1})^{-1} &&= A - A(A+B)^{-1}A = A(A+B)^{-1}B \\ \mu &= \boldsymbol{\Sigma}(A^{-1}\mu_1 + B^{-1}\mu_2) &&= \mu_1 + A(A+B)^{-1}(\mu_2 - \mu_1) \end{aligned}$$

according to the following table:

	A	μ_1	B	μ_2
$\boldsymbol{y}_\star^M \mid \bar{y}^F, \boldsymbol{y}^M, \boldsymbol{b}$	$\boldsymbol{\Sigma}^F$	$\bar{y}^F - \boldsymbol{b}$	$\boldsymbol{\Sigma}_{\star\mid\bullet}$	$\mu_{\star\mid\bullet}$
$\boldsymbol{b} \mid \bar{y}^F, \boldsymbol{y}^M, \boldsymbol{y}_\star^M$	$\boldsymbol{\Sigma}^F$	$\bar{y}^F - \boldsymbol{y}_\star^M$	$\boldsymbol{\Sigma}^b$	μ^b
$\boldsymbol{y}_\star^M \mid \bar{y}^F, \boldsymbol{y}^M$	$\boldsymbol{\Sigma}_{\star\mid\bullet}$	$\mu_{\star\mid\bullet}$	$\boldsymbol{\Sigma}^b + \boldsymbol{\Sigma}^F$	$\bar{y}^F - \mu^b$
$\boldsymbol{b} \mid \bar{y}^F, \boldsymbol{y}^M$	$\boldsymbol{\Sigma}^F + \boldsymbol{\Sigma}_{\star\mid\bullet}$	$y^F - \mu_{\star\mid\bullet}$	$\boldsymbol{\Sigma}^b$	μ^b

4 Notes

The quantity θ^M never seems to be used (although we sometimes condition on it). In context the functions $c^b(\cdot,\cdot)$ and $c^M(\cdot,\cdot)$ must be correlation functions (probably from the power-exponential family) for the discrepancy/bias and for the model, respectively. Perhaps θ^M includes the range and power parameters for $c^M(\cdot,\cdot)$? What about $c^b(\cdot,\cdot)$?

In (4) the prior mean of $\boldsymbol{b}$ is $\boldsymbol{\mu}^b$, but $\mu^b = 0$ is assumed in Section (2). For example, the conditional distribution of λ^b (item 3) would have rate parameter $r_{b|\bullet} = r_b + (\boldsymbol{b} - \mu^b)'[c^b(D^F, D^F)]^{-1}(\boldsymbol{b} - \mu^b)/2$ (I think) if μ^b doesn't vanish.

Last edited by RLW: February 18, 2014, 11:38 EST