Introducing multivator: A Multivariate Emulator

Robin K. S. Hankin

Auckland University of Technology

Abstract

A multivariate generalization of the **emulator** technique described by Hankin (2005) is presented in which random multivariate functions may be assessed. In the standard univariate case (Oakley 1999), a Gaussian process, a finite number of observations is made; here, observations of different types are considered. The technique has the property that marginal analysis (that is, considering only a single observation type) reduces exactly to the univariate theory.

The associated software is used to analyze datasets from the field of climate change.

This vignette is based on Hankin (2012).

1. Introduction

Many scientific disciplines require the use of complex computer models. Such models, also known as "simulators", are valid objects of inference and are often assumed to be random functions and assessed using the Bayesian statistical paradigm (Currin, Mitchell, Morris, and Ylvisaker 1991); in particular, computer models are often assumed to be Gaussian Processes (Oakley and O'Hagan 2002).

Although deterministic—in the sense that running the simulator twice with identical inputs gives identical outputs—the Bayesian paradigm is to treat the code output as a random variable because, before the computational task is finished, one has subjective uncertainty about the outcome; de Finetti (1974) discusses the philosophy of this approach. Hankin (2005) discusses this issue from a computational perspective.

Given that the simulator is a random function, uncertainty about its behaviour is reducible to an arbitrarily low level by running the simulator sufficiently many times. However, because many modern simulators require large amounts of computer time to run, this is not possible; in practice one is typically presented with a fixed number of simulator runs as data.

One tool used to make inferences about simulators under these circumstances is the *emulator* (Oakley 1999), and the **BACCO** suite of R packages (Hankin 2005). The emulator is an established technique that has been used in many fields including Earth systems science (McNeall 2008), oceanography (Challenor, Hankin, and Marsh 2006), and climate science (Warren *et al.* 2008). However, **BACCO** is limited to univariate random functions. In this paper, I present a generalization of the Gaussian Process which allows the technique to be used for multivariate simulator output.

The current work is a generalization of that of Conti and O'Hagan (2010), who presented a

separable covariance structure. Here, I present a generalization of that work in which the roughness lengths of the components of the multivariate process are allowed to differ.

1.1. Review of theory for the univariate emulator

This section presents a very brief review of the univariate emulator. Much of the material is taken directly from Oakley (1999) with slight changes of notation.

For any random univariate function $\eta \colon \mathbb{R}^d \to \mathbb{R}$ and set of points $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ with $\mathbf{x}_i \in \mathbb{R}^d$, the random vector $\mathbf{y} = (\eta(\mathbf{x}_1), \ldots, \eta(\mathbf{x}_n))^\top$ is assumed to be multivariate normal:

$$\mathbf{y} | \boldsymbol{\beta}, \Sigma \sim \mathcal{N}\left(H\boldsymbol{\beta}, \Sigma\right) \tag{1}$$

where $H = (\mathbf{h}(\mathbf{x}_1), \ldots, \mathbf{h}(\mathbf{x}_n))^\top$ is the matrix of (known) regressor functions $\mathbf{h} \colon \mathbb{R}^d \to \mathbb{R}^q$ so the regressor matrix H is n by q, denoted $H_{[n \times q]}$; it is sometimes convenient to write $H = H(\mathbf{X})$ where $\mathbf{X}_{[n \times d]}$ is the design matrix. Equation 1 is conditional on the (unknown) vector of coefficients $\boldsymbol{\beta}_{[q]}$ and the variance matrix $\Sigma_{[n \times n]}$. A common choice for $\mathbf{h}(\cdot)$ is $\mathbf{h}(\mathbf{x}) = (1, x_1, \ldots, x_d)^\top$ [thus $q = d + 1$], but one is in principle free to choose any function of $\mathbf{x}$.

The variance matrix is, explicitly:

$$\Sigma = \begin{pmatrix} \mathsf{VAR}(\eta(\mathbf{x}_1)) & \mathsf{COV}(\eta(\mathbf{x}_1), \eta(\mathbf{x}_2)) & \cdots & \mathsf{COV}(\eta(\mathbf{x}_1), \eta(\mathbf{x}_n)) \\ \mathsf{COV}(\eta(\mathbf{x}_2), \eta(\mathbf{x}_1)) & \mathsf{VAR}(\eta(\mathbf{x}_2)) & & \vdots \\ \vdots & & \ddots & \\ \mathsf{COV}(\eta(\mathbf{x}_n), \eta(\mathbf{x}_1)) & \cdots & & \mathsf{VAR}(\eta(\mathbf{x}_n)) \end{pmatrix}. \tag{2}$$

(Oakley writes $\sigma^2 A$ for Σ, where $A_{[n \times n]}$ is a matrix of correlations and σ^2 is an overall variance). It can be shown that

$$\eta(\cdot) | \boldsymbol{\beta}, \Sigma, \mathbf{y} \sim \mathcal{N}\left(m^*(\cdot), \mathsf{COV}^*(\cdot, \cdot)\right) \tag{3}$$

where

$$
\begin{aligned}
m^*(\mathbf{x}) &= \mathbf{h}(\mathbf{x})^\top \boldsymbol{\beta} + \mathbf{t}(\mathbf{x})^\top \Sigma^{-1} \left(\mathbf{y} - H\boldsymbol{\beta}\right) & (4) \\
\mathsf{COV}^*(\eta(\mathbf{x}), \eta(\mathbf{x}')) &= \mathsf{COV}(\eta(\mathbf{x}), \eta(\mathbf{x}')) - \mathbf{t}(\mathbf{x})^\top \Sigma^{-1} \mathbf{t}(\mathbf{x}') & (5) \\
\mathbf{t}(\mathbf{x})^\top &= \left(\mathsf{COV}(\eta(\mathbf{x}), \eta(\mathbf{x}_1)), \ldots, \mathsf{COV}(\eta(\mathbf{x}), \eta(\mathbf{x}_n))\right) & (6) \\
\mathbf{y}^\top &= \left(\eta(\mathbf{x}_1), \ldots, \eta(\mathbf{x}_n)\right). & (7)
\end{aligned}
$$

If an improper flat prior for $\boldsymbol{\beta}$ is used, its posterior conditional distribution can be shown to be

$$\boldsymbol{\beta} | \Sigma, \mathbf{y} \sim \mathcal{N}\left(\hat{\boldsymbol{\beta}}, \left(H^\top \Sigma^{-1} H\right)^{-1}\right)$$

where

$$\hat{\boldsymbol{\beta}} = \left(H^\top \Sigma^{-1} H\right)^{-1} H^\top \Sigma^{-1} \mathbf{y}$$

is the posterior mean. It is possible to integrate out $\boldsymbol{\beta}$ to obtain

$$\eta(\cdot) | \Sigma \sim \mathcal{N}\left(m^{**}(\cdot), \mathsf{COV}^{**}(\cdot, \cdot)\right) \tag{8}$$

where

$$m^{**}(\mathbf{x}) \;=\; \mathbf{h}(\mathbf{x})^{\top}\hat{\boldsymbol{\beta}} + \mathbf{t}(\mathbf{x})^{\top}\Sigma^{-1}\left(\mathbf{y} - H\hat{\boldsymbol{\beta}}\right) \tag{9}$$

$$\mathsf{COV}^{**}(\eta(\mathbf{x}),\eta(\mathbf{x}')) \;=\; \mathsf{COV}^{*}(\eta(\mathbf{x}),\eta(\mathbf{x}')) +$$
$$\left(\mathbf{h}(\mathbf{x})^{\top} - \mathbf{t}(\mathbf{x})^{\top}\Sigma^{-1}H\right)\left(H^{\top}\Sigma^{-1}H\right)^{-1}\left(\mathbf{h}(\mathbf{x}')^{\top} - \mathbf{t}(\mathbf{x}')^{\top}\Sigma^{-1}H\right)^{\top} \tag{10}$$

The two superscript stars mean that the results have been integrated with respect to the posterior distribution of $\boldsymbol{\beta}$. What these equations mean is that

$$\eta(\cdot)|\,\Sigma \sim \mathcal{N}\left(m^{**}(\cdot), \mathsf{COV}^{**}\left(\eta(\cdot),\eta(\cdot)\right)\right). \tag{11}$$

Or, in words, that $m^{**}(\mathbf{x})$ is a quick approximation for the $\eta(\mathbf{x})$ in the sense that its posterior distribution is Gaussian with mean and variance given by the right hand side of Equation 9 and 10 respectively. It is usual to refer to Equation 11 as the *emulator*; observe that the entire posterior distribution is specified.

Positive definiteness

The covariance matrix, Equation 2, is required to be positive definite for any choice of design matrix. This can be guaranteed by appropriate choice of covariance function.

Writing $\mathsf{COV}(\eta(\mathbf{x}),\eta(\mathbf{x}')) = \sigma^2 c(\mathbf{x} - \mathbf{x}')$, then Bochner's theorem (Feller 1966) shows that Σ is positive definite for any $\mathbf{x}_1,\ldots,\mathbf{x}_n$ if and only if $c(\mathbf{t})$ is the characteristic function of a symmetric probability Borel measure:

$$c(\mathbf{t}) = \int_{\boldsymbol{\omega}\in\mathbb{R}^d} e^{i\boldsymbol{\omega}^{\top}\mathbf{t}}\,dF(\boldsymbol{\omega}). \tag{12}$$

One standard choice (Hankin 2005) is a standard multivariate Gaussian distribution[1] with mean zero and variance $\mathcal{S}_{[d\times d]}$. This gives

$$c(\mathbf{t}) = \int_{\boldsymbol{\omega}\in\mathbb{R}^d} e^{i\boldsymbol{\omega}^{\top}\mathbf{t}}\frac{1}{\sqrt{|2\pi\mathcal{S}|}}\exp(-\tfrac{1}{2}\boldsymbol{\omega}^{\top}\mathcal{S}^{-1}\boldsymbol{\omega})\,d\boldsymbol{\omega}. \tag{13}$$

In practice one writes $\mathcal{B} = \mathcal{S}^{-1}/2$ and absorbs the normalization constant into a σ^2 term leaving:

$$c(\mathbf{t}) = \exp\{-\mathbf{t}^{\top}\mathcal{B}\mathbf{t}\} \tag{14}$$

giving

$$\mathsf{COV}(\eta(\mathbf{x}),\eta(\mathbf{x}')) = \sigma^2 c(\mathbf{x}-\mathbf{x}') = \sigma^2\exp\{-\mathbf{t}^{\top}\mathcal{B}\mathbf{t}\}. \tag{15}$$

Then Σ in equation 2 is guaranteed to be positive-definite.

---

[1] A number of different choices for $f(\cdot)$ have been used in the literature. Stein (1999), for example, advocates a Student t- distribution, but the corresponding generalization of Equation 13 is the subject of "controversy and difficulties" (Dreier and Kotz 2002), possessing no closed form solution (Sutradhar 1986), and further work would be needed to implement it in the context of **BACCO**.

2. Earlier multivariate work

A natural generalization of Equation 3 is to consider $\boldsymbol{\eta}\colon \mathbb{R}^d \to \mathbb{R}^p$ with a separable covariance function. Conti and O'Hagan (2010), for example, generalized equation 8 to a matrix Gaussian with a column covariance matrix given by Equation 15, and a row covariance matrix $\Lambda_{[p \times p]}$ which they treated as an additional hyperparameter.

Rougier (2008), considering the common problem of multidimensional model output that is indexed by a Cartesian grid, presented a computationally advantageous method; and Higdon (2008) considered the principal components of multivariate experimental results.

However, all these approaches suffer from the disadvantage that the separability of the covariance matrix implies that the roughness lengths of each of the components are identical. This assumption is often not justified: For example, in climatology, although rainfall and temperature are correlated, orographic effects mean that spatial correlation lengths are smaller for rainfall than temperature. The simple example given in Section 5.1 uses terminology inspired by this motivating example.

2.1. Non-separable covariance structures

To accommodate differing roughness lengths, it is necessary to use non-separable covariance structures. Examples include that of Majumdar and Gelfand (2007), who observed that the convolution of two positive-definite covariance functions is again positive definite. However, Majumdar and Gelfand noted that in practice the convolution will have no closed form, a drawback not affecting the present work.

Recent unpublished work by Fricker, Oakley, and Urban (2010) also uses convolution techniques and presents a nonseparable covariance structure of which the present work is shown to be a generalization.

Related work might also include Kennedy and O'Hagan (2000)[2], who presented methods to analyze a hierarchy of levels of a model. The present work, however, does not make the Markov assumption (their Equation 1), and does not have the nested design restriction ($D_t \subseteq D_{t-1}$ in their notation).

2.2. Dimension reduction and Bayesian estimation

Highly multivariate output (such as a temperature field over a 3D Cartesian lattice) is difficult to deal with and many workers have sought methods to reduce such output to a more manageable format. The techniques discussed above are a special case of dimension reduction but other techniques have been presented in the context of Bayesian inference.

Principal Component Analysis is one frequently used tool. Higdon (2008), for example, considers high dimensional data from a series of experiments involving high explosive and applies the methods of Kennedy and O'Hagan (2001), although the principal components are assumed to be independent, an assumption not necessary in the present approach.

Other techniques for dimension reduction exist. Bayarri, Berger, Cafeo, Garcia-Donato, Liu, Palomo, Parthasarathy, Paulo, Sacks, and Walsh (2007), for example, use wavelet decomposition and use a thresholding procedure to produce a manageable number of coefficients. The

---

[2]The **approximator** package (Hankin 2009) provides a suite of related R functionality.

techniques outlined in the present paper are applicable in principle to wavelet decompositions, but further work would be needed.

3. The multivariate case

In this section, I outline a scheme by which the emulator of Section 1.1 may be generalized to the multivariate case in a computationally tractable manner, with exact expressions for the (conditional) covariance matrix 2. The presentation uses a generalization of Bochner's theorem in such a way as to precisely delineate the space of admissible covariance functions.

In the multivariate case, there are p different types of observation, say $\eta_r(\cdot)$ for $r = 1, \ldots, p$. Each type of observation is a Gaussian process (hence susceptible to analysis by the **emulator** package), but here we admit covariances between the observation types, so that $\mathsf{COV}(\eta_r(\mathbf{x}), \eta_s(\mathbf{x}')) \neq 0$ for $r \neq s$. Here $\eta_r(\mathbf{x})$ is the value of an observation of type r at point $\mathbf{x}$.

We suppose that observations of type r are made at points $\mathbf{X}^{(r)} = \left(\mathbf{X}_1^{(r)}, \ldots, \mathbf{X}_{n_r}^{(r)}\right)^\top$ for $1 \leqslant r \leqslant p$. Thus observations of type r are made at points on a design matrix $\mathbf{X}_{[n_r \times d]}^{(r)}$.

It is straightforward to specify the expectation. This is just

$$
\mathsf{E}(\mathbf{d}) = H\boldsymbol{\beta} = \begin{pmatrix} \mathbf{h}_1\left(\mathbf{X}^{(1)}\right)^\top & 0 & \cdots & 0 \\ 0 & \mathbf{h}_2\left(\mathbf{X}^{(2)}\right)^\top & & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & & \mathbf{h}_p\left(\mathbf{X}^{(p)}\right)^\top \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1 \\ \vdots \\ \boldsymbol{\beta}_p \end{pmatrix} \tag{16}
$$

where $\mathbf{h}_r(\cdot)$ are the basis functions for the observation types r with $1 \leqslant r \leqslant p$; thus $H_{[\sum_{r=1}^p n_r \times \sum_{r=1}^p q_r]}$ is a generalized regressor matrix. See how the overall coefficient vector $\boldsymbol{\beta} = \left(\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_p\right)^\top$ may be partitioned into its several components. It is not necessary for all the $\boldsymbol{\beta}_r$ to be of the same length.

The overall variance matrix will be

$$
\Sigma = \begin{bmatrix} \Sigma^{(11)} & \Sigma^{(12)} & \cdots & \Sigma^{(1p)} \\ \Sigma^{(21)} & \Sigma^{(22)} & \cdots & \Sigma^{(2p)} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma^{(p1)} & \Sigma^{(p2)} & \cdots & \Sigma^{(pp)} \end{bmatrix} \tag{17}
$$

where $\Sigma^{(rs)}$ refers to the covariance between observations of type r and s, specifically $\Sigma^{(rr)}$ are the restricted univariate variance matrices for observation type $r = 1, \ldots, p$ and the off-diagonal entries represent covariances.

Generalization to the multivariate case is subtle. We seek a method of determining Σ of Equation 17 in such as way that $\Sigma^{(rr)}$ may be specified using standard techniques (typically from a univariate analysis; the $\Sigma^{(rr)}$ being determined on the basis of different $\mathcal{B}_r$ in 15 in general), *and* the $\Sigma^{(rs)}, r \neq s$ represent covariances between observations of type r and s in a reasonable way. It is necessary to guarantee that Σ in Equation 17 is positive definite.

Formally, we seek functions $C_{rs}(\cdot, \cdot)$ with $C_{rs}(\mathbf{x}, \mathbf{x}') = \mathsf{COV}(\eta_r(\mathbf{x}), \eta_s(\mathbf{x}'))$. In the notation of Equation 17, we would have $\Sigma^{(rs)} = C_{rs}\left(\mathbf{X}^{(r)}, \mathbf{X}^{(s)}\right)$ as the matrix of covariances between

observations of type r at $\mathbf{X}^{(r)}$ and observations of type s at $\mathbf{X}^{(s)}$, that is, between $\eta_r\left(\mathbf{X}^{(r)}\right)$ and $\eta_s\left(\mathbf{X}^{(s)}\right)$. These functions are required to be positive-definite in the sense that Σ of Equation 17 must be positive definite for any set of points $\mathbf{X}^{(1)},\ldots,\mathbf{X}^{(p)}$.

A matrix generalization of 12 was presented by Cramer (1940) which will be used here: $C_{rs}(\mathbf{t})$ are positive-definite if and only if they are of the form

$$C_{rs}(\mathbf{t}) = \int_{\boldsymbol{\omega}\in\mathbb{R}^q} e^{i\boldsymbol{\omega}^\top\cdot\mathbf{t}} dF_{rs}(\boldsymbol{\omega}) \tag{18}$$

for some positive definite $F_{ij}(\boldsymbol{\omega})$. If attention is restricted to absolutely integrable functions (a condition which will be dropped subsequently), this becomes

$$C_{rs}(\mathbf{t}) = \int_{\boldsymbol{\omega}\in\mathbb{R}^q} e^{i\boldsymbol{\omega}^\top\cdot\mathbf{t}} f_{rs}(\boldsymbol{\omega})\, d\boldsymbol{\omega}. \tag{19}$$

If we write $||f(\boldsymbol{\omega})||$ for the matrix with (r,s) entry $f_{rs}(\boldsymbol{\omega})$, then we require $||f(\boldsymbol{\omega})||_{[p\times p]}$ to be positive definite for all $\boldsymbol{\omega}$.

Considering functions of the form discussed in Equation 13, one approach would be to specify the off-diagonal elements to be zero. Here $p = 3$ is used for illustration; the general case follows directly:

$$||f(\boldsymbol{\omega})|| = \begin{bmatrix} \dfrac{\exp\{-\frac{1}{2}\boldsymbol{\omega}^\top\mathcal{S}_1^{-1}\boldsymbol{\omega}\}}{|2\pi\mathcal{S}_1|^{1/2}} & 0 & 0 \\ 0 & \dfrac{\exp\{-\frac{1}{2}\boldsymbol{\omega}^\top\mathcal{S}_2^{-1}\boldsymbol{\omega}\}}{|2\pi\mathcal{S}_2|^{1/2}} & 0 \\ 0 & 0 & \dfrac{\exp\{-\frac{1}{2}\boldsymbol{\omega}^\top\mathcal{S}_3^{-1}\boldsymbol{\omega}\}}{|2\pi\mathcal{S}_3|^{1/2}} \end{bmatrix} \tag{20}$$

where the $\mathcal{S}_i$ are positive-definite matrices corresponding to the (marginal) univariate covariance matrices of Equation 15. This matrix is positive definite for all $\boldsymbol{\omega}$. This approach would only be appropriate if the covariances between observation types were zero.

One way to account for nonzero covariance between observation types is suggested by the fact that, given positive numbers $x_1,\ldots,x_p$, the matrix with element (r,s) equal to $\sqrt{x_r x_s}$ is positive semidefinite. Thus

$$||f(\boldsymbol{\omega})|| = \begin{bmatrix} \dfrac{\exp\{-\frac{1}{2}\boldsymbol{\omega}^\top\mathcal{S}_1^{-1}\boldsymbol{\omega}\}}{|2\pi\mathcal{S}_1|^{1/2}} & \dfrac{\exp\{-\frac{1}{2}\boldsymbol{\omega}^\top(\frac{1}{2}\mathcal{S}_1^{-1}+\frac{1}{2}\mathcal{S}_2^{-1})\boldsymbol{\omega}\}}{|2\pi\mathcal{S}_1|^{1/4}\cdot|2\pi\mathcal{S}_2|^{1/4}} & \dfrac{\exp\{-\frac{1}{2}\boldsymbol{\omega}^\top(\frac{1}{2}\mathcal{S}_1^{-1}+\frac{1}{2}\mathcal{S}_3^{-1})\boldsymbol{\omega}\}}{|2\pi\mathcal{S}_1|^{1/4}\cdot|2\pi\mathcal{S}_3|^{1/4}} \\ \dfrac{\exp\{-\frac{1}{2}\boldsymbol{\omega}^\top(\frac{1}{2}\mathcal{S}_2^{-1}+\frac{1}{2}\mathcal{S}_1^{-1})\boldsymbol{\omega}\}}{|2\pi\mathcal{S}_2|^{1/4}\cdot|2\pi\mathcal{S}_1|^{1/4}} & \dfrac{\exp\{-\frac{1}{2}\boldsymbol{\omega}^\top\mathcal{S}_2^{-1}\boldsymbol{\omega}\}}{|2\pi\mathcal{S}_2|^{1/2}} & \dfrac{\exp\{-\frac{1}{2}\boldsymbol{\omega}^\top(\frac{1}{2}\mathcal{S}_2^{-1}+\frac{1}{2}\mathcal{S}_3^{-1})\boldsymbol{\omega}\}}{|2\pi\mathcal{S}_2|^{1/4}\cdot|2\pi\mathcal{S}_3|^{1/4}} \\ \dfrac{\exp\{-\frac{1}{2}\boldsymbol{\omega}^\top(\frac{1}{2}\mathcal{S}_3^{-1}+\frac{1}{2}\mathcal{S}_1^{-1})\boldsymbol{\omega}\}}{|2\pi\mathcal{S}_3|^{1/4}\cdot|2\pi\mathcal{S}_1|^{1/4}} & \dfrac{\exp\{-\frac{1}{2}\boldsymbol{\omega}^\top(\frac{1}{2}\mathcal{S}_3^{-1}+\frac{1}{2}\mathcal{S}_2^{-1})\boldsymbol{\omega}\}}{|2\pi\mathcal{S}_3|^{1/4}\cdot|2\pi\mathcal{S}_2|^{1/4}} & \dfrac{\exp\{-\frac{1}{2}\boldsymbol{\omega}^\top\mathcal{S}_3^{-1}\boldsymbol{\omega}\}}{|2\pi\mathcal{S}_3|^{1/2}} \end{bmatrix} \tag{21}$$

is positive semidefinite for all $\boldsymbol{\omega}$ provided only that the $\mathcal{S}_i$ are positive definite; observe that the diagonal elements of Equations 20 and 21 match. Observe that, with fixed diagonal entries, offdiagonal elements cannot exceed those given in Equation 21 while retaining positive definiteness. The matrix thus corresponds to "maximal correlation" in this sense and the general terms are then:

$$C_{rs}(\mathbf{t}) = \begin{cases} \exp\left\{-\mathbf{t}^\top \mathcal{B}_r \mathbf{t}\right\} & \text{if } r = s \\ \dfrac{\exp\left\{-\mathbf{t}^\top \left(\frac{1}{2}\mathcal{B}_r^{-1} + \frac{1}{2}\mathcal{B}_s^{-1}\right)^{-1}\mathbf{t}\right\}}{\left|\left(\frac{1}{2}\mathcal{B}_r + \frac{1}{2}\mathcal{B}_s\right)\left(\frac{1}{2}\mathcal{B}_s^{-1} + \frac{1}{2}\mathcal{B}_2^{-1}\right)\right|^{1/4}} & \text{otherwise} \end{cases} \tag{22}$$

where we follow standard convention (Oakley 1999) and write $\mathcal{B}_r = \mathcal{S}_r^{-1}/2$. Similar expressions occur in the study of nonstationary covariance functions (Paciorek and Schervish 2006; Higdon 2002); a special case (diagonal matrices) is given by Fricker *et al.* (2010). These authors construct the covariance matrix using process convolutions, observing that the convolution theorem for Fourier transforms ensures positive definiteness (Higdon 2002, 2008).

Equation 22 gives a positive-semidefinite variance matrix for any design matrix. Noting that the Schur (elementwise) product of a positive-semidefinite matrix and a positive definite matrix is positive definite, the relation

$$C'_{rs}(\mathbf{t}) = M_{rs}C_{rs}(\mathbf{t}) \tag{23}$$

is a positive definite function. Here $M_{[p\times p]}$ is a positive-definite matrix that accounts for covariance between observation types.

Other forms for the covariance matrix

It is possible to use covariance functions other than the Gaussian form used in Equation 21. The probability measures are required to be symmetric, and the geometric mean of two measures is required to have a characteristic function in closed form.

Measures that are proportional to an indicator function, that is

$$I_A(\mathbf{x}) = \begin{cases} C & \text{if } \mathbf{x} \in A \\ 0 & \text{otherwise} \end{cases}$$

where C is the normalization constant and $A \subset \mathbb{R}^d$ is a support set, are a natural choice. In this case element (i, j) would be $I_{A_i \cap A_j}(\mathbf{x})$; one could consider support sets that are hyperspheres or, more interestingly, orthotopes.

One other natural choice would be the multivariate t-distribution, but further work would be necessary to assess its suitability in this context.

Summary

The univariate emulator is generalized to the p- variate case. Univariate expectation $H\boldsymbol{\beta}$ is generalized to the multivariate form given in Equation 16, and the univariate variance matrix of Equation 2 is generalized to the multivariate form 17 with

$$\left[\Sigma^{(rs)}\right]_{ij} = M_{rs}\frac{\exp\left\{-\left(\mathbf{x}_i^{(r)} - \mathbf{x}_j^{(s)}\right)^\top \left(\frac{1}{2}\mathcal{B}_r^{-1} + \frac{1}{2}\mathcal{B}_s^{-1}\right)^{-1}\left(\mathbf{x}_i^{(r)} - \mathbf{x}_j^{(s)}\right)\right\}}{\left|\left(\frac{1}{2}\mathcal{B}_r + \frac{1}{2}\mathcal{B}_s\right)\left(\frac{1}{2}\mathcal{B}_r^{-1} + \frac{1}{2}\mathcal{B}_s^{-1}\right)\right|^{1/4}}. \tag{24}$$

arising from the positive definite function $C'(\cdot, \cdot)$ of Equation 23. The matrices $\Sigma^{(rr)}$ correspond to univariate variance matrices and each is obtained from a matrix $\mathcal{B}_i$ of roughnesses in
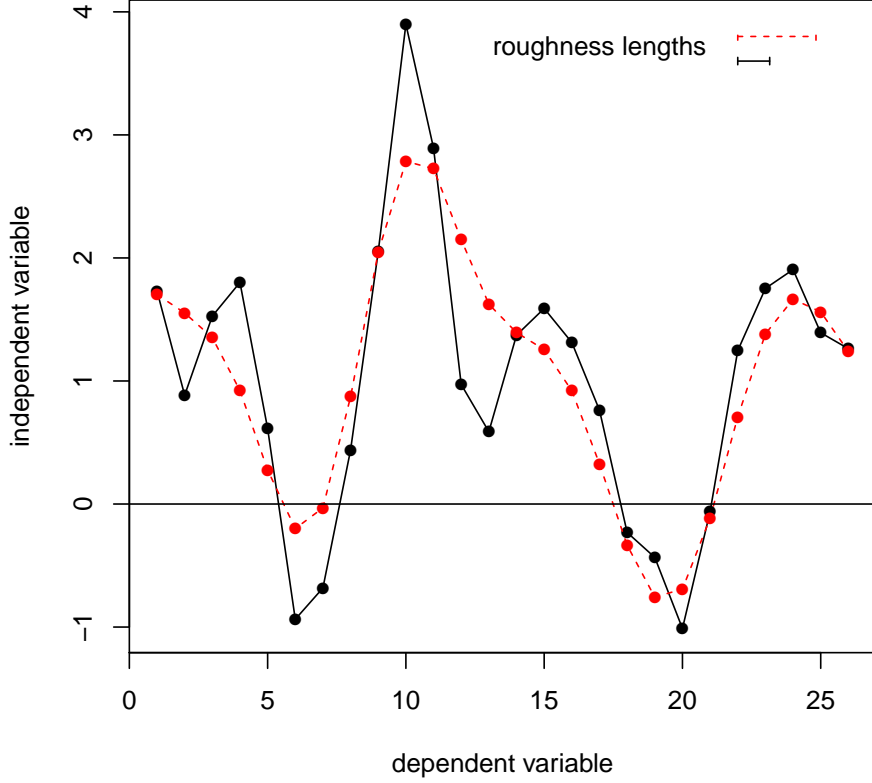
Figure 1: An example of two correlated Gaussian processes with different roughness lengths, indicated on the diagram. See how the red curve, having a longer roughness length, is smoother than the black curve with a shorter roughness length.

the same way as in the univariate case. The univariate variance σ^2 generalizes to a variance matrix M whose diagonal elements correspond to the univariate variances $\sigma_i^2, 1 \leqslant i \leqslant p$.

3.1. Discussion

The above analysis suggests a method by which a covariance matrix may be determined for multivariate observations. Here I discuss some implications of Equation 24.

Consider the case where $\mathcal{B}_i$ are known. Then consider the correlation $c(\cdot, \cdot)$ between the types of observations at the same point, ie $\mathbf{t} = \mathbf{0}$ in Equation 22. This is just

$$|c(r,s)| \leqslant \left| \left(\tfrac{1}{2}\mathcal{B}_r + \tfrac{1}{2}\mathcal{B}_s \right) \left(\tfrac{1}{2}\mathcal{B}_r^{-1} + \tfrac{1}{2}\mathcal{B}_s^{-1} \right) \right|^{-1/4} \leqslant 1 \tag{25}$$

where the first inequality is sharp if and only if $M_{rs}^2 = M_{rr}M_{ss}$; observe that M being positive semidefinite implies $M_{rs}^2 \leqslant M_{rr}M_{ss}$. The second inequality follows from the concavity of $\log |\mathcal{B}|$ (Cover and Thomas 1988) and is thus sharp if and only if $\mathcal{B}_r = \mathcal{B}_s$. In

the case of 1×1 matrices (ie scalars), the matrices commute and the maximum correlation is $\left[\frac{1}{2}\left(1 + \mathcal{B}_r \mathcal{B}_s^{-1} + \mathcal{B}_s \mathcal{B}_r^{-1}\right)\right]^{-1/4}$ (Figure 1 shows an example of two maximally correlated Gaussian processes with different roughness lengths).

Thus $\mathcal{B}_r \neq \mathcal{B}_s$ imposes an active upper bound on $c(r, s)$: two Gaussian processes with different roughness coefficients cannot be perfectly correlated.

It is also evident that

$$\mathsf{COV}\left(\eta_r\left(\mathbf{x}_1\right), \eta_s\left(\mathbf{x}_2\right)\right) = \mathsf{COV}\left(\eta_r\left(\mathbf{x}_2\right), \eta_s\left(\mathbf{x}_1\right)\right), \tag{26}$$

for any matrices $M, \mathcal{B}_r, \mathcal{B}_s$.

4. Estimation of hyperparameters

The **multivator** package requires a generalized set of hyperparameters compared with the **emulator** package. The **emulator** package needs a single positive-definite matrix $\mathcal{B}$ that expressed the roughness length of the response function; **multivator** requires matrices $\mathcal{B}_1, \ldots, \mathcal{B}_p$: One matrix per type of observation. Each matrix represents the marginal roughness characteristics of each observation type.

Oakley (1999), and many subsequent authors, assumed that the overall variance matrix Σ was given by $\Sigma = \sigma^2 A$, where σ^2 is a scalar and A a matrix of correlations. Oakley (1999) proceeded to integrate out σ^2 (using a weak prior distribution) to obtain an expression for the posterior distribution of the process in terms of $\hat{\sigma}^2$, the estimated value for σ^2.

The approach advocated here, by contrast, generalizes the scalar variance σ^2 to M, a $p \times p$ positive-definite matrix which expresses the overall variances and covariances of the p different types of observation; subsequent analysis is conditional on the values of the $\mathcal{B}_i$ and M.

The procedure used in the package is a three step process:

1. Estimate the roughness parameters for each observation type separately, using techniques of the **emulator** package,

2. Calculate the marginal variance terms, using an analytical expression for the posterior mode, following Oakley (1999); these are the diagonal elements of M,

3. Estimate the off-diagonal elements of M by numerical determination of the posterior mode. To ensure positive-definiteness, an improper flat prior with nonzero support extending over the positive-definite matrices may be used.

This multi-stage procedure is reminiscent of the two-stage process outlined in Kennedy and O'Hagan (2001). It seems to work reasonably well in practice. The process is not perfect: One might wish to calculate the joint likelihood of M and the $\mathcal{B}_i$ simultaneously; the relevant likelihood is given by Oakley's Equation 2.36, which in our notation is

$$\mathcal{L}\left(M, \mathcal{B}_1, \ldots, \mathcal{B}_p\right) = \frac{\left|\Sigma^{-1}\right|^{1/2}}{\left|H^\top \Sigma^{-1} H\right|^{1/2}} \exp\left(-\frac{1}{2}\left(d - H\hat{\boldsymbol{\beta}}\right)^\top \Sigma^{-1}\left(d - H\hat{\boldsymbol{\beta}}\right)\right), \tag{27}$$

and optimize that, but such an approach seems impractical, even for the toy example considered here.

5. The package in use

The **multivator** package of R (R Development Core Team 2011) routines is now demonstrated using three examples: A toy dataset in which the underlying assumptions are *known* to be true; evaluates of a simple function, following Oakley (1999); and a larger dataset drawn from the discipline of physical oceanography. A brief discussion of the package as applied to modular systems such as CIAS (Warren *et al.* 2008) is also given.

5.1. Toy example

Although the toy dataset and associated R objects are simple, they represent the most general form of the package's functionality and furnish a comprehensive suite of tests of the package functionality.

Toy dataset `toy_mm` is a simple design matrix on three levels: `temp`, `rain`, and `humidity`.

```
R> data("mtoys")
R> head(toy_mm)


       a      b     c      d     type
t1 0.500 0.5000 0.500 0.5000     temp
r1 0.500 0.5000 0.500 0.5000     rain
h1 0.500 0.5000 0.500 0.5000 humidity
t2 0.017 0.0966 0.812 0.4148     temp
t3 0.938 0.5284 0.619 0.2102     temp
t4 0.983 0.3239 0.472 0.0966     temp
```

Thus `toy_mm` is a multivariate design matrix, typically a latin hypercube. Observations on `toy_mm` are provided in `toy_d`:

```
R> head(toy_d)

  t1   r1   h1   t2   t3   t4
8.60 5.70 3.10 8.75 8.01 6.56
```

The central function of the package is `multem()`, corresponding to `interp()` of package **emulator**. Suppose we wish to make inferences about a particular point in parameter space:

```
R> toy_point

     a   b   c   d     type
t1 0.3 0.4 0.5 0.6     temp
r1 0.3 0.4 0.5 0.6     rain
h1 0.3 0.4 0.5 0.6 humidity
```

Thus `toy_point` corresponds to measuring all three levels (`temp`, `rain`, `humidity`) at a single point in parameter space. It is straightforward to use the package to provide an estimate for the process at this point, using `multem()`:

```
R> (e <- multem(toy_point, toy_expt, toy_mhp, toy_LoF, give = TRUE))
```

```
$mstar
  t1   r1   h1
8.64 5.41 2.41
```

```
$cstar
        t1     r1     h1
t1   0.2700 -0.139 0.0577
r1 -0.1390  0.717 0.3057
h1  0.0577  0.306 0.6435
```

[Object `toy_expt` is an S4 object with slots for the design matrix and observations, produced by `experiment()`]. The return value of `multem()` is a two-element list with the first being a vector whose elements are the posterior mean for each row of the multivariate design matrix `toy_mm`, and the second is the conditional variance matrix. Thus we see that, at this point in parameter space, temperature and rainfall are negatively correlated. The diagonal of the matrix gives the (conditional) marginal variances for the three levels (`temp`, `rain`, `humidity`).

So, for example, one might sample from the posterior distribution:

```
R> rmvnorm(n=5,mean=e$mstar,sigma=e$cstar)
```

```
       t1   r1   h1
[1,] 9.40 5.26 3.45
[2,] 9.14 5.30 1.38
[3,] 8.21 5.27 2.29
[4,] 9.71 5.60 2.09
[5,] 8.08 5.25 2.05
```

The equivalent univariate analysis may be carried out using function `interpolant.quick()` of the **emulator** package:

```
R> interpolant.quick(
+                    x      = x_uni,
+                    d      = d_uni,
+                    xold   = m_uni,
+                    Ainv   = A_uni,
+                    scales = s_uni,
+                    func   = f_uni,
+                    give.Z = TRUE)
```

```
$mstar.star
[1] 8.76
```

```
$Z
[1] 0.622
```

```
$prior
[1] 8.54
```

[the `_uni` suffix denotes the univariate subset corresponding to `temp`]. The mean value changes from about 8.64 in the multivariate case to 8.76 in the univariate case, and the conditional variance changes from about 0.27 to about $0.622^2 = 0.387$. The difference is due to the nonindependence of the observation types.

5.2. Estimation of the hyperparameters in the package

In this section, the hyperparameters for the synthetic dataset considered above are estimated using the package, following the scheme suggested above.

In common with the **emulator** and **calibrator** packages, the **multivator** package includes functionality to create datasets with values drawn from the appropriate distribution.

```
R> mm <- toy_mm_maker(81,82,83)
R> d <- obs_maker(mm, toy_mhp, toy_LoF, toy_beta)
R> jj_expt <- experiment(mm,d)
```

Here `mm` is a multivariate design matrix, created using a latin hypercube; the three arguments specify the number of points in parameter space at which each observation type is made. Function `obs_maker()` creates observations drawn from the appropriate distribution. Here, `toy_mhp` is a hyperparameter object (a matrix $M_{[3\times3]}$ of covariances, and three $\mathcal{B}_{[4\times4]}$ roughness matrices, one per observation type); `toy_LoF` is a list of regressor functions, and `toy_beta` is a vector of regression coefficients.

The function `optimal_scales()` first estimates the $\mathcal{B}_i$ matrices and then, conditional on this, estimates the overall covariance matrix M, conditional on the $\mathcal{B}_i$, using Equation 27:

```
R> mhp_opt <- optimal_params(jj_expt, toy_LoF, option="b")
```

Specifying `option="b"` restricts the $\mathcal{B}_i$ to diagonal matrices. The optimized value for M, the matrix of covariances is then given by

```
R> M(mhp_opt)
```

```
            temp    rain humidity
temp       1.552  -1.354    0.286
rain      -1.354   1.345    0.299
humidity   0.286   0.299    1.898
```

Compare the true value:

```
R> M(toy_mhp)
```

```
         temp rain humidity
temp      1.0 -0.7      0.5
rain     -0.7  2.0      1.4
humidity  0.5  1.4      3.0
```

5.3. Validation

It is possible to validate the above approach by the technique of using half the dataset for fitting the emulator (as above), then the remaining half for validation. The appropriate R expression would be

```
R> est2 <- multem(toy_mm2, toy_expt, toy_mhp, toy_LoF)
```

where `toy_mm` and `toy_mm2` are components of the *same* multivariate observation taken from the distribution specified in Equation 1. Figure 2 shows such an exercise, exhibiting reasonable agreement between observed and predicated values.

6. Simple functional analysis

In this section, a simple function $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ is considered, and univariate inference is compared with the multivariate techniques introduced above.

From a computational perspective, an analysis using the **multivator** package is presented "from scratch"; standard R objects are coerced to the appropriate S4 objects.

The functions considered are $f_a(x, y) = \sin(5 \cdot (x+y))$ and $f_b(x, y) = 7\sin(5 \cdot (x+y)) + \sin(20 \cdot (x - y))$. These functions correspond to observations of type 'a' and 'b' respectively, and are chosen so that they are correlated, but f_a might be expected to have a smoother response than f_b. An experimental design is then needed for each function, which in this case is a simple latin hypercube:

```
R> # number of observation points:
R> na <- 33    # observation of 'a'
R> nb <- 09    # observation of 'b'

R> xa <- latin.hypercube(na,2) # so rows of 'xa' are observation points for 'a'
R> xb <- xa[seq_len(nb),]
R> #xb <- latin.hypercube(nb,2)
```

Thus `xa` and `xb` are standard R matrices. It is now possible to evaluate f_a and f_b over their experimental designs:

```
R> a_obs <- apply(xa,1,fa)
R> b_obs <- apply(xb,1,fb)
```

Thus there are two design matrices `xa` and `xb`, and two corresponding sets of observations, here `a_obs` and `b_obs`, all in the form of standard R objects (matrices and vectors respectively). It is now straightforward to apply the **multivator** package methods.
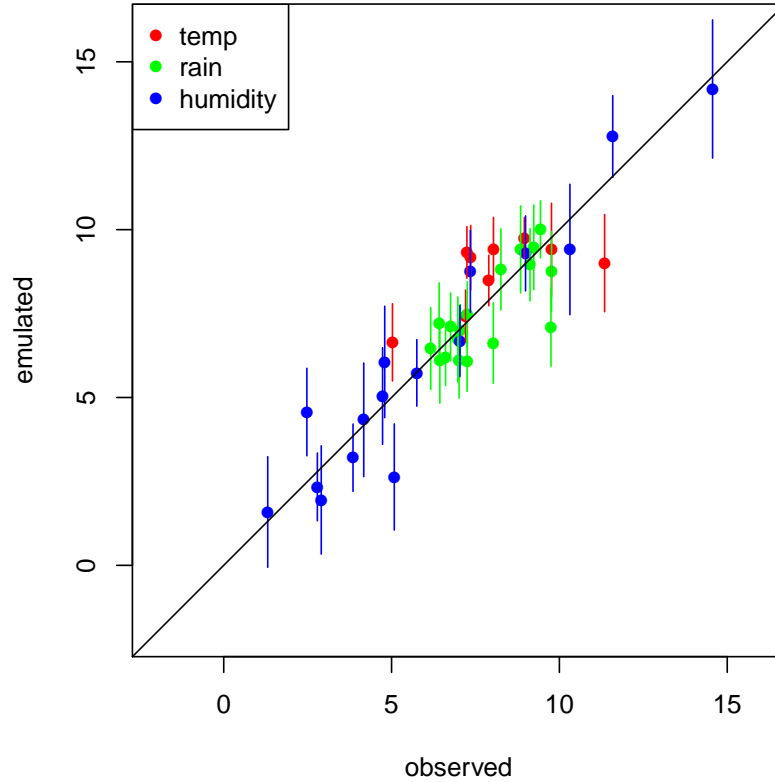
Figure 2: Observed vs predicted values for a sample from the multivariate Gaussian distribution defined by Equation 1 with mean and variance defined by Equations 16 and 24. Error bars correspond to marginal standard deviations.

We first define a multivariate design matrix (an object of class "`mdm`") by combining the univariate design matrices `xa` and `xb`, then create an `experiment` object by adding the code observations; and finally estimate optimal parameters using `optimal_params()`:

```
R> RS_mdm <- mdm(rbind(xa,xb),types=c(rep("a",na),rep("b",nb)))
R> RS_expt <- experiment(mm=RS_mdm, obs= c(a_obs,b_obs))
R> RS_opt <- optimal_params(RS_expt, option="b")
```

The three objects above define a working multivariate emulator in terms of bespoke S4 objects specific to the **multivator** package. Suppose we wish to predict f_b and f_b on a set of $n = 20$ points in its domain:

```
R> n <- 20
R> xnew <- latin.hypercube(n,2,names=c("x","y"))
R> #xnew <- cbind(x=runif(20),y=runif(20))
```

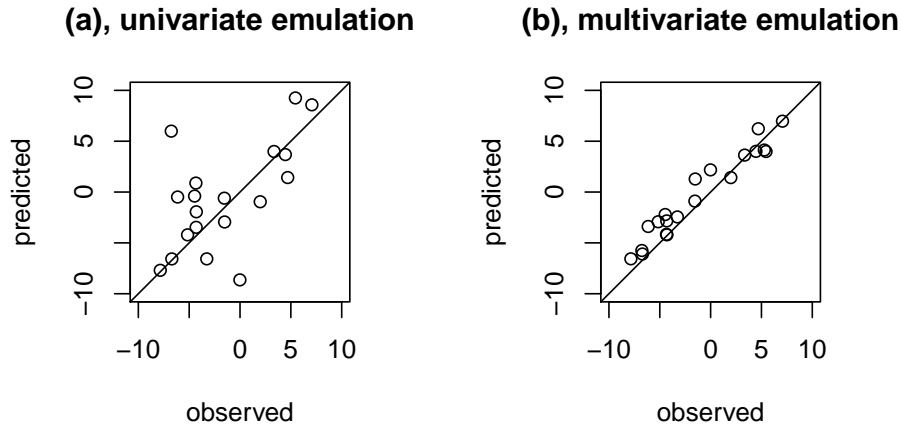**(a), univariate emulation**   **(b), multivariate emulation**

Figure 3: Analysis of a simple function on a 2D Latin hypercube. Function evaluations on the horizontal axis plotted against predicted values on the vertical axis. (a), univariate emulation ($R^2 = 0.429$) and (b), multivariate emulation ($R^2 = 0.942$).

The appropriate R idiom is to create a new multivariate design matrix on `xnew`; then use function `multem()` to provide multivariate estimates of f_b on the design matrix:

```
R> RS_new_mdm <- mdm(rbind(xnew,xnew),rep(c("a","b"),each=n))
R> RS_prediction <- multem(x=RS_new_mdm, expt=RS_expt, hp=RS_opt)
```

A graphical summary of the results is given in Figure 3.

7. Data analysis using multivator

The package is now used to analyze climate change data obtained from the Genie-Goldstein model, a computationally efficient Earth-Systems model designed to assess climate change from an oceanographical perspective on a timescale of centuries to millennia (Edwards and Marsh 2005).
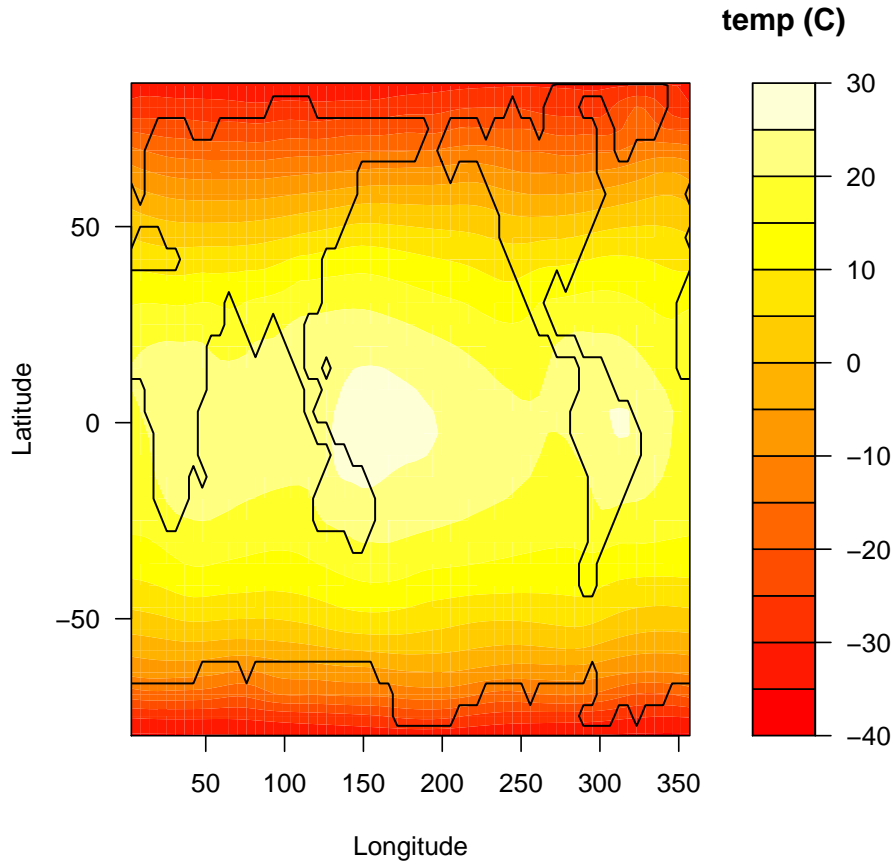
Figure 4: Typical output from Genie-Goldstein: A global map of temperature, interpreted as yearly average values.

McNeall (2008) considered Genie-Goldstein output and used Principal Component Analysis as an analytic technique. Here, I consider the first four principal components using the **multivator** package. Although principal components are mutually orthogonal, they are not necessarily independent with respect to any given regressors. I now show how data provided by McNeall may be analyzed using the `multivator` package.

```
R> data("mcneall")
R> dim(mcneall_temps)
```

```
[1] 2048    92
```

The `mcneall_temps` matrix has 92 columns, one for each of 92 runs of Genie-Goldstein. Each column, of 2048 numbers, corresponds to a map of global temperature; an example is given in Figure 4 in which the `showmap()` function is used to reshape the vector to a form suitable for display.

Dataset `mcneall_pc` has 92 rows, one per run, and 20 columns. The first 16 columns show the design matrix of independent variables[3]. The last four columns are the first four principal components of the output; an interpretation is given in Figure 5.

```
R> dim(mcneall_pc)

[1] 92 20

R> head(mcneall_pc,2)

        WSF    AHD    AMD  WAHDP  ZHAF   MHAF   ZMAF    MMAF   CRF   THP
[1,] 0.310 0.3774 0.924 0.3515 0.894 0.449 0.791 0.0988 0.216 0.11
[2,] 0.278 0.0963 0.451 0.0951 0.209 0.364 0.203 0.5285 0.825 0.30
       APMF    SC   OHD   OVD   ODC   SID     pc1      pc2     pc3
[1,] 0.340 0.528 0.401 0.672 0.590 0.435 -0.0335 -0.01571  0.0545
[2,] 0.472 0.087 0.682 0.510 0.435 0.730 -0.2129  0.00541 -0.1584
        pc4
[1,] 0.0398
[2,] 0.0782
```

Although this dataset is more involved than the others considered in this paper, the same computational techniques may be used:

```
R> jj <- apart(mcneall_pc, 17:20)
R> opt_mcneall <- optimal_params(jj, start_hp=opt_mcneall, option='a')
```

Then we may examine the covariance matrix between residuals of the first four principal components:

```
R> (CM <- M(opt_mcneall))

          pc1       pc2       pc3       pc4
pc1  6.08e-04 -0.000149 -5.63e-05 -0.000674
pc2 -1.49e-04  0.001954 -1.74e-04  0.000396
pc3 -5.63e-05 -0.000174  5.12e-03  0.001365
pc4 -6.74e-04  0.000396  1.36e-03  0.007301
```

This shows that the correlations between the principal components are nontrivial:

```
R> CM/sqrt(tcrossprod(diag(CM)))

        pc1    pc2     pc3     pc4
pc1  1.0000 -0.137 -0.0319 -0.320
pc2 -0.1369  1.000 -0.0550  0.105
pc3 -0.0319 -0.055  1.0000  0.223
pc4 -0.3201  0.105  0.2233  1.000
```

---

[3]That is, physical parameters with uncertain values, needed as inputs to Genie-Goldstein; the first one, 'WSF', for example, is 'windstress scaling factor'; McNeall gives a full discussion and a table on page 50.

Principal component 1

Principal component 3

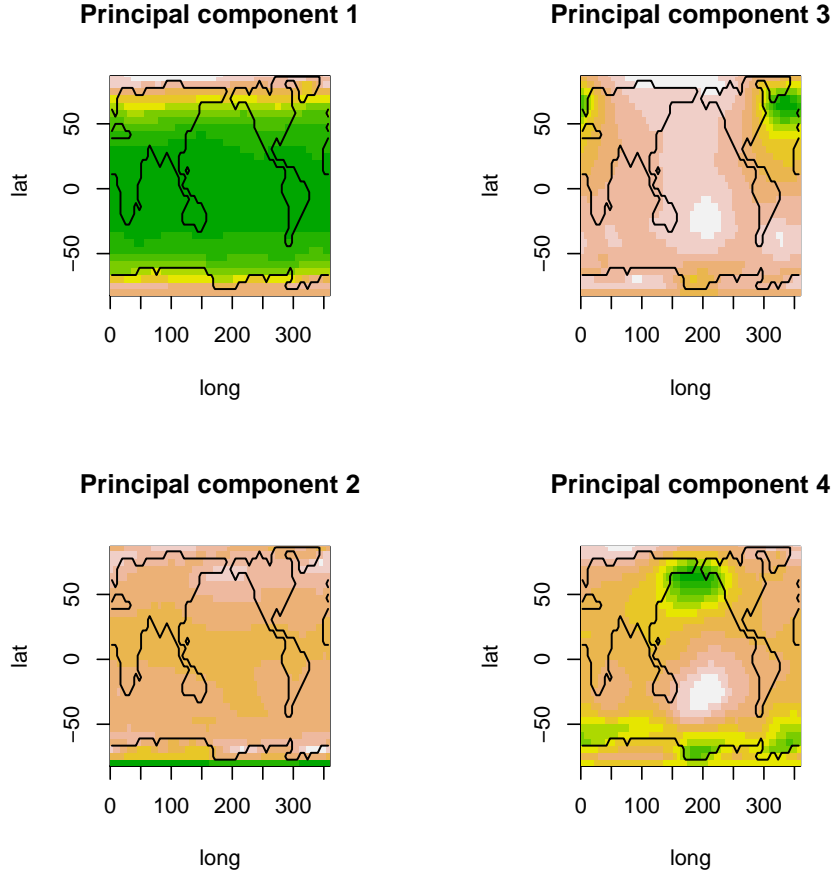**Principal component 2**

Principal component 4

Figure 5: The first four principal components in the McNeall dataset of 92 Genie-Goldstein runs. The first shows the standard Pole/Equator variability; the second shows the uncertainty near the South Pole; the third represents uncertainty due to the bistability of the meridional overturning circulation in the North Atlantic; the fourth appears to be related to the state of the Pacific Decadal Oscillation or possibly the ENSO.

In particular, the positive correlation between the third and fourth component may be interpreted from the perspective of more sophisticated approaches such as general circulation models (Wan 2009).

Note that these correlations are conditional upon the form of the regressor functions (here the default set `default_LoF`).

7.1. Modular systems

Multivariate emulation appears to be a useful technique in the context of modular systems such as CIAS (Warren *et al.* 2008) in which a model comprises various component "modules".

In the case of CIAS, the modules address various aspects of the global climate system and examples include E3MG which models the global economy, MAGICC which models the phys-

ical global climate, and ICLIPS which models the impacts of climate change. The modules exchange information at runtime using the BFG protocol.

One feature of CIAS is that it is possible to replace any module with another functionally equivalent one. Multivariate emulation is useful when considering the behaviour of CIAS used in this way. If one has p different interchangeable modules, then the output of CIAS is a p-variate random variable that may be analyzed using the **multivator** package.

In an associated vignette, visible from within an R session by typing `vignette("cias")`, a short analysis of a synthetic dataset is presented.

8. Discussion

A generalization of the emulator to multivariate datasets is proposed and the **multivator** package has been introduced. The package is used to analyze datasets drawn from the fields of oceanography and climate change. The variance structure proposed appears to have pleasing and useful properties. Further work might include extension of the ideas presented here to complex functions.

References

Bayarri MJ, Berger JO, Cafeo J, Garcia-Donato G, Liu F, Palomo J, Parthasarathy RJ, Paulo R, Sacks J, Walsh D (2007). "Computer Model Validation with Functional Output." *The Annals of Statistics*, **35**(5), 1874–1906.

Challenor P, Hankin RKS, Marsh R (2006). *Avoiding Dangerous Climate Change*, chapter 7, "Towards the Probability of Rapid Climate Change", pp. 55–63. Cambridge University Press.

Conti S, O'Hagan A (2010). "Bayesian Emulation of Complex Multi-Output and Dynamic Computer Models." *Journal of Statistical Planning and Inference*, **140**, 640–651.

Cover TM, Thomas JA (1988). "Determinant Inequalities Via Information Theory." *Siam Journal on Matrix Analysis and Applications*, **9**(3), 384–392.

Cramer H (1940). "On the Theory of Stationary Random Processes." *The Annals of Mathematics*, **41**(1), 215–230. ISSN 0003486X. URL http://www.jstor.org/stable/1968827.

Currin C, Mitchell TJ, Morris M, Ylvisaker D (1991). "Bayesian Prediction of Deterministic Functions with Applications to the Design and Analysis of Computer Experiments." *Journal of the American Statistical Association*, **86**, 953–963.

de Finetti B (1974). *Theory of Probability*. John Wiley & Sons.

Dreier I, Kotz S (2002). "A Note on the Characteristic Function of the t-Distribution." *Statistics and Probability Letters*, **57**, 221–224.

Edwards N, Marsh R (2005). "Uncertainties due to Transport-Parameter Sensitivity in an Efficient 3D Ocean-Climate Model." *Climate Dynamics*, **24**, 415–433.

Feller W (1966). *An Introduction to Probability Theory and its Applications*, volume two. John Wiley & Sons.

Fricker T, Oakley J, Urban NM (2010). "Multivariate Emulators with Nonseparable Covariance Structures." URL http://www.mucm.ac.uk/Pages/Dissemination/Dissemination_Papers_Technical.html.

Hankin RKS (2005). "Introducing **BACCO**, an R Bundle for Bayesian Analysis of Computer Code Output." *Journal of Statistical Software*, **14**(16).

Hankin RKS (2009). *approximator: Bayesian Prediction Of Complex Computer Codes*. R package version 1.1-6, URL http://CRAN.R-project.org/package=approximator.

Hankin RKS (2012). "Introducing **multivator**: A Multivariate Emulator." *Journal of Statistical Software*, **46**(8), 1–20. URL http://www.jstatsoft.org/v46/i08/.

Higdon D (2002). "Space and Space-Time Modeling Using Process Convolutions." In CW Anderson, V Barnett, PC Chatwin, AH El-Shaarawi (eds.), *Quantitative Methods for Current Environmental Issues*, pp. 37–56. Springer-Verlag. ISBN 1852332948.

Higdon D (2008). "Computer Model Calibration Using High-Dimensional Output." *Journal of the American Statistical Association*, **103**(482), 570–583.

Kennedy MC, O'Hagan A (2000). "Predicting the Output From a Complex Computer Code When Fast Approximations Are Available." *Biometrika*, **87**(1), 1–13.

Kennedy MC, O'Hagan A (2001). "Bayesian Calibration Of Computer Models." *Journal of the Royal Statistical Society B*, **63**(3), 425–464.

Majumdar A, Gelfand AE (2007). "Multivariate Spatial Modeling for Geostatistical Data Using Convolved Covariance Functions." *Mathematical Geology*, **39**(2), 225–245.

McNeall DJ (2008). *Dimension Reduction in the Bayesian Analysis of a Numerical Climate Model*. Ph.D. thesis, University of Southampton.

Oakley J (1999). *Bayesian Uncertainty Analysis For Complex Computer Codes*. Ph.D. thesis, University of Sheffield.

Oakley J, O'Hagan A (2002). "Bayesian Inference For The Uncertainty Distribution Of Computer Model Outputs." *Biometrika*, **89**(4), 769–784.

Paciorek CJ, Schervish MJ (2006). "Spatial Modelling Using a New Class of Nonstationary Covariance Functions." *Environmetrics*, **17**(5), 483–506. doi:10.1002/env.785. URL http://dx.doi.org/10.1002/env.785.

R Development Core Team (2011). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Rougier J (2008). "Efficient Emulators for Multivariate Deterministic Functions." *Journal of Computational and Graphical Statistics*, **17**(4), 827–843.

Stein ML (1999). *Interpolation of Spatial Data: Some Theory for Kriging.* Springer-Verlag.

Sutradhar BC (1986). "On the Characteristic Function of Multivariate Student t-Distribution." *The Canadian Journal of Statistics*, **14**(4), 329–337.

Wan X (2009). *Effect of Atlantic Meridional Overturning Circulation Changes on Tropical Coupled Ocean-Atmosphere System.* Ph.D. thesis, Texas A&M University.

Warren R, *et al.* (2008). "Development and Illustrative Outputs Of The Community Integrated Assessment System (CIAS), a Multi-Institutional Modular Integrated Assessment Approach for Modelling Climate Change." *Environmental Modelling and Software*, **23**, 592–610.

Affiliation:

Robin K. S. Hankin
Auckland University of Technology
School of Computing and Mathematical Sciences
Wakefield Street
Auckland
New Zealand
E-mail: hankin.robin@gmail.com