J Am Stat Assoc. Author manuscript; available in PMC 2016 Apr 18.
Published in final edited form as:
J Am Stat Assoc. 2012; 107(500): 1395–1409.
Published online 2012 Jul 30. doi: 10.1080/01621459.2012.712414
PMCID: PMC4834988
NIHMSID: NIHMS751571
PMID: 27099406

Minkowski–Weyl Priors for Models With Parameter Constraints: An Analysis of the BioCycle Study

Michelle R. Danaher, Anindya Roy, Zhen Chen, Sunni L. Mumford, and Enrique F. Schisterman

Michelle R. Danaher, Predoctoral Fellow, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, Rockville, MD 20852;
Contributor Information.
Michelle R. Danaher: danahermr@mail.nih.gov; Anindya Roy: anindya@umbc.edu; Zhen Chen: chenzhe@mail.nih.gov; Sunni L. Mumford: mumfords@mail.nih.gov; Enrique F. Schisterman: schistee@mail.nih.gov

Abstract

We propose a general framework for performing full Bayesian analysis under linear inequality parameter constraints. The proposal is motivated by the BioCycle Study, a large cohort study of hormone levels of healthy women where certain well-established linear inequality constraints on the log-hormone levels should be accounted for in the statistical inferential procedure. Based on the Minkowski–Weyl decomposition of polyhedral regions, we propose a class of priors that are fully supported on the parameter space with linear inequality constraints, and we fit a Bayesian linear mixed model to the BioCycle data using such a prior. We observe positive associations between estrogen and progesterone levels and F_2-isoprostanes, a marker for oxidative stress. These findings are of particular interest to reproductive epidemiologists.

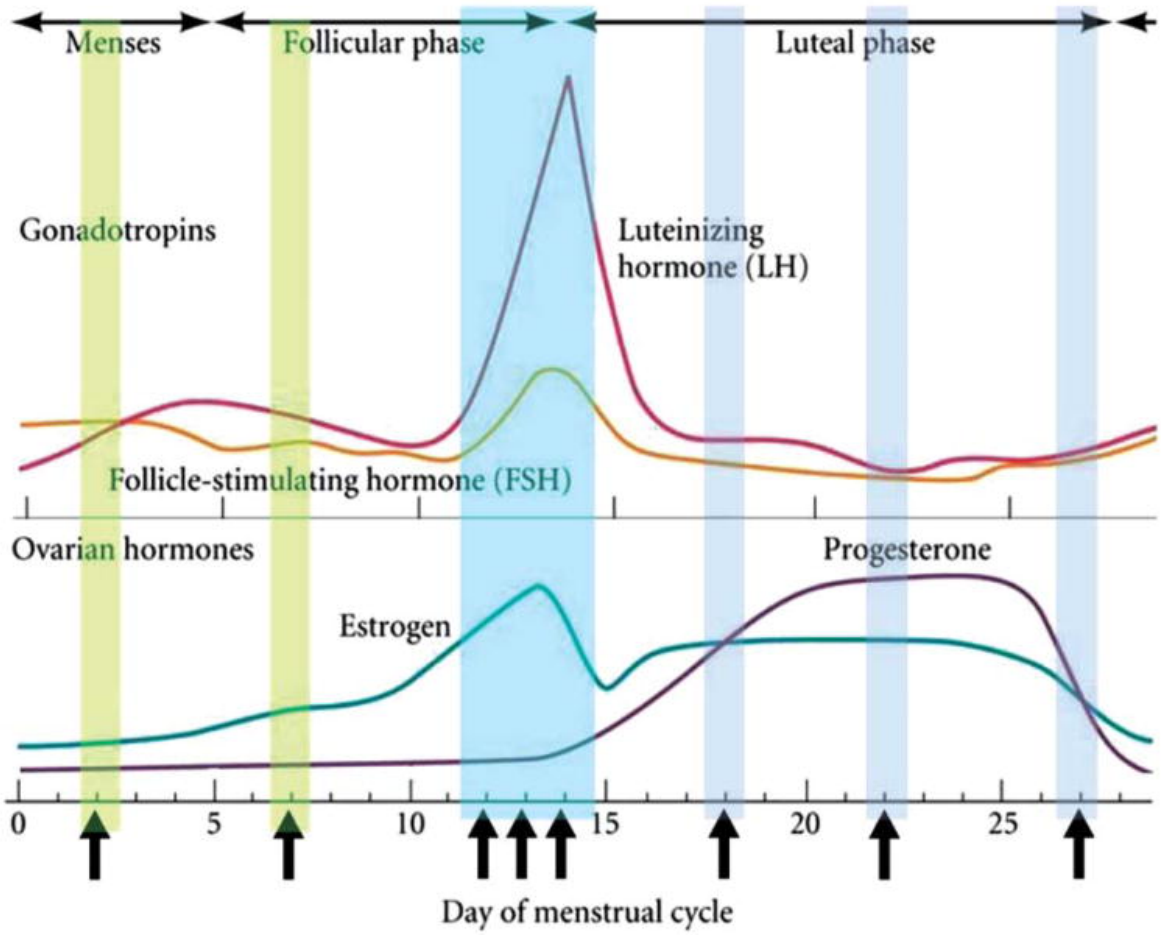
Keywords: Bayesian inference, Extreme directions, Extreme points, Parameter restriction, Polyhedral region

1. INTRODUCTION

Changes in hormone levels during the menstrual cycle are regulated by complex feedback mechanisms (Hall 2009). Understanding the sources of variation in hormonal patterns is of great interest to reproductive epidemiologists, as hormone levels have been associated with various chronic diseases and breast and endometrial cancers (Fritz and Speroff 1982). Moreover, changes in hormonal patterns may be associated with, or indicative of, other biochemical activity such as rapid lipid peroxidation.

The BioCycle Study is a relatively large cohort study with a unique study design that was conducted by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development for the purpose of understanding reproductive hormonal patterns and their association with other biological processes during the menstrual cycle among healthy, regularly menstruating women (Wactawski-Wende et al. 2009). While the BioCycle Study presents many biological and statistical modeling challenges with its unique sampling design and its wealth of data, the primary motivation for this article addresses efficiently handling complex parameter constraints that are naturally imposed by the well-understood physiology of hormonal dynamics during the menstrual cycles of healthy women.

To describe the main problem being addressed here, we need first to understand the BioCycle Study design. The BioCycle Study was a prospective cohort of 259 healthy, regularly menstruating women aged 18–44 years from western New York State (Wactawski-Wende et al. 2009). To participate, women had to report regular cycles for the past 6 months (between 21 and 35 days). Women with conditions known to affect menstrual cycle function such as polycystic ovary syndrome, uterine fibroids, or current use of hormonal contraception (i.e., 3 months prior to study entry) were excluded. Women were provided with fertility monitors to schedule clinic visits during phases of the menstrual cycle with the most hormonal variability (Howards et al. 2009). Eight clinic visits were scheduled per cycle: one visit during menses, one visit in the mid-follicular phase, three visits around expected ovulation, and three visits in the luteal phase. Fasting serum samples were collected at each clinic visit for measurement of hormone levels (including estrogen, luteinizing hormone (LH), follicle stimulating hormone (FSH), and progesterone) and other biological markers. Figure 1 illustrates where the eight visits were expected to sample corresponding to predicted hormone profiles (Gilbert 2006). Due to the study design, the hormone trajectories are essentially indexed by the different phases of a regular menstrual cycle. The hormone levels at different phases may be compared qualitatively based on existing knowledge. One such qualitative constraint is obtained from the LH surge that triggers ovulation. Levels of LH during ovulation (visit 4) are expected to be much higher (at least 50%) than those during the follicular phase (visit 2). Since the hormones are usually modeled/analyzed on the log scale, this qualitative constraint translates into a linear constraint involving the mean LH levels during visits 2 and 4. Once all such plausible constraints among the levels of the four reproductive hormones are put together, it constitutes a complicated polygonal constraint set involving the 32 parameters associated with the mean levels of the four hormones at the eight clinic visits during the menstrual cycle. Accounting for these linear constraints among the parameters will likely improve efficiency and obtain biologically plausible interpretations of the associations between hormonal patterns and other factors.



Open in a separate window

Figure 1

Predicted hormonal profiles. The curves have been scaled to represent them on the same plot. Upward arrows represent expected timing of eight scheduled clinic visits relative to menstrual cycle phase. Figure adapted from Gilbert (2006). The online version of this figure is in color.

There is a long history of incorporating known parameter constraints in inferential statistics procedures to improve efficiency. Books by Robertson, Wright, and Dykstra (1988) and Silvapulle and Sen (2004) provide numerous examples of, mainly frequentist, procedures under parameter restrictions. The restricted maximum likelihood estimator provides an extension of the maximum likelihood estimator to models with parameter restrictions. In comparison, the literature on Bayesian inference under parameter constraints appears somewhat limited. This may be because the Bayesian formulation comes more naturally and may appear too straightforward to merit further research. The Bayesian paradigm allows the information about parameter restrictions to be perspicuously used in the model building exercise via prior specification and thereby often provides a more apt representation of the actual

physical system. Still, the vigorous debate about the merits and shortcomings of both frequentist and Bayesian methods when applied to models with parameter restrictions goes on, see Mandelkern (2002) and the comments in the discussion.

In many situations with parameter constraints, a Bayesian solution is known to have favorable properties, for example, Marchand and MacGibbon (2000); Marchand and Perron (2001). In spite of the criticism of subjectivity of the prior specification, Bayesian solutions often have desirable risk properties. The advantage of the Bayesian solution is more apparent in the constrained parameter situation where ad hoc frequentist procedures may become inadmissible. As argued eloquently by van Dyk (2002), the Bayesian solution is far less subjective in spite of the task of subjective prior specification. A Bayesian solution with a weakly informative prior supported on the constraint space will be reasonably objective while being an admissible solution.

One of the challenges in obtaining a Bayesian solution is specification of a prior distribution that is supported on the constraint space. This is particularly difficult when the complexity of the constraint space is high compared to the dimension. Our primary interest in the context of the BioCycle Study will be to address the issue of parameter constraints in a linear mixed model with linear inequality constraints on some of the fixed effect parameters. Davis (1978), Geweke (1986, 1996), and Chen and Deely (1996) looked at Bayesian inference problems for regression parameters with linear inequality constraints. Linear inequality constraints appear routinely in social sciences, including behaviorial science (Klugkist, Laudy, and Hoijtink 2005b), economics (Pindyck and Rubinfeld 1981), physics (Mandelkern 2002), and of course in biological sciences, which is the context of the present article. Gelman, Meng, and Stern (1996) discussed the effect of inequality constraints in the context of posterior predictive model assessment, and Klugkist, Laudy, and Hoijtink (2005a,b); Stern (2005); Kato and Hoijtink (2006); Mulder et al. (2009); and Mulder, Hoijtink, and Klugkist (2010) provided in-depth discussions of the effect of linear constraints on Bayesian model selection.

Gelfand, Smith, and Lee (1992) gave a general recommendation for performing Bayesian inference under parameter restrictions. Their approach recommends obtaining posterior draws under an unrestricted prior and collecting only those draws that satisfy the constraints. Thus, for any problem where it is feasible to generate posterior samples under an unrestricted prior, it is possible to generate samples when the prior is truncated to the constraint space. Chen and Shao (1998) discussed posterior computation for such an approach. The appeal of this method is its simplicity. However, when the constraint space has negligible mass under the unrestricted prior, this method can be very computationally inefficient. For complex parametric models with a large number of parameters and constraints, the truncation method may not be practical.

To remedy this situation, approximate Bayesian schemes have been proposed in many applications. Hoffman (1993) used a generalized Bayesian solution for regression problems with linear inequality constraints. Gunn and Dunson (2005) used a transformation approach which is an attractive alternative because it provides computational efficiency while truncating the prior to the closure of the constraint space. This approach draws efficient inference in shape restricted regression problems. Roy et al. (2011) generalized the transformation approach to include general linear inequality constraints. These approaches are computationally more efficient than the truncation approach, especially when the constraint space has low posterior mass under an unrestricted prior. However, the transformation

approach tends to accumulate mass on the boundary of the constraint space, a feature that may not be desirable in some applications. Indeed authors have discussed scenarios where it is desirable to have priors with no mass on the boundary to avoid situations where the estimate falls on the boundary of the constraint space (Galindo-Garre and Vermunt 2006; Chung et al. 2011). From a decision theoretic point of view, somewhat subjective inference with respect to a reasonably diffused but fully supported prior may be preferred over a prior with mass on the boundary. In this article, we suggest a flexible class of priors that are fully supported on a constraint space defined by general linear inequality constraints, thereby providing a general method for obtaining a full Bayesian solution for problems with linear inequality constraints.

We motivate and describe our methodology in the setting of a linear (mixed) model because of its appropriateness and prior usage as a model for analyzing the data from the BioCycle Study. In Section 2, we use a simple linear model for description and illustration of the proposed prior and provide some numerical results. The chosen model is a simplification of the model for the BioCycle data that is given in Section 4. The simple model is chosen mainly for the convenience of pictorial depictions of the constraints and the posterior which is not possible in more complex higher dimensional settings. We provide a detailed analysis of the BioCycle data using the proposed prior in Section 4. The Bayesian proposal that we describe for models with linear parametric restrictions is applicable for any model with linear inequality constraints on parameters.

2. MINKOWSKI–WEYL PRIOR

For illustration, consider a Gaussian linear model where it is known that the mean parameter $\boldsymbol{\mu}$ satisfies linear inequality constraints. Specifically, let $\boldsymbol{y}_i = \boldsymbol{\mu} + \boldsymbol{\varepsilon}_i$, where $\boldsymbol{\varepsilon}_i \overset{\text{iid}}{\sim} N_p(\mathbf{0}, \boldsymbol{\Sigma})$ and $\boldsymbol{y}_i$ is the $p \times 1$ vector of the ith observation. Suppose there is prior knowledge that the mean parameter, $\boldsymbol{\mu}$, belongs to the convex polyhedral set

$$\mathcal{M} = \{\boldsymbol{\mu} \in \mathbb{R}^p : \boldsymbol{A}\boldsymbol{\mu} < \boldsymbol{b}, \boldsymbol{\mu} > 0\}, \tag{1}$$

where $\boldsymbol{A}$ is a $c \times p$ matrix and $\boldsymbol{b}$ is a $c \times 1$ vector. The likelihood is then given by

$$\mathcal{L}(\boldsymbol{\mu}, \boldsymbol{\Sigma} \mid \boldsymbol{y}_1, \ldots, \boldsymbol{y}_n) = \prod_{i=1}^{n} N_p(\boldsymbol{y}_i; \boldsymbol{\mu}, \boldsymbol{\Sigma}) I(\boldsymbol{\mu} \in \mathcal{M}). \tag{2}$$

The covariance parameter $\boldsymbol{\Sigma}$ is taken to belong to $\mathbb{C}$, the cone of all $p \times p$ positive definite matrices. Bayesian analysis would require one to specify a prior on the restricted set $\mathcal{M} \times \mathbb{C}$. While there are standard recommendations for specifying a marginal prior on $\mathbb{C}$, specifying a prior on $\mathcal{M}$ may be challenging. A natural way of invoking a prior on $\mathcal{M}$ would be to truncate an unrestricted prior on $\boldsymbol{\mu}$ to the set $\mathcal{M}$. The prior density will be the product of the density of the unrestricted prior and the indicator function $I(\boldsymbol{\mu} \in \mathcal{M})$. Geweke (1986) used an importance sampling technique to generate posterior samples from such a prior. Gelfand, Smith, and Lee (1992) made the truncation method more widely applicable by proposing the following scheme: use a Markov chain Monte Carlo (MCMC) method to

draw a posterior sample using an unrestricted prior on $\boldsymbol{\mu}$ and discard the draws that do not belong to $\mathscr{M}$. While these methods have the appeal of simplicity, computationally they can be quite inefficient, especially if the prior mass on the constraint set is very small. Specification of priors that are fully supported on the restricted set $\mathscr{M}$ can be quite difficult if the parameter dimension is moderately large and the polyhedral region described by the constraints is complicated.

We propose a new class of priors that directly take into account the geometry of the constraint set and provides a flexible specification of priors for $\boldsymbol{\mu}$ that are fully supported on $\mathscr{M}$. The prior specification relies on a key geometric decomposition of the polyhedral set in terms of a simplex and a cone, both of which allow greater flexibility in terms of prior specification than the polyhedral region. For convenience, we restate the geometry result of the Minkowski–Weyl (MW) representation theorem for polyhedral regions.

[Minkowski-Weyl decomposition]: *A set* $\wp$ *in* $\mathbb{R}^p$ *is a convex polyhedral iff*

$$\mathscr{P} = \text{conv}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k) \oplus \text{cone}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_l)$$

for some vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_l$ *in* $\mathbb{R}^p$ *where* $\oplus$ *denotes the standard direct sum for sets*.

The region conv($\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$) represents the convex hull of the vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ and the region cone($\boldsymbol{u}_1, \ldots, \boldsymbol{u}_l$) represents the cone defined by the linear combination of the vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_l$ with positive coefficients. The vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ denote the k extreme points of the polyhedral $\wp$ and the vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_l$ represent the extreme directions. Thus, any vector $\boldsymbol{\theta}$ in $\wp$ can be represented as

$$\boldsymbol{\theta} = (\alpha_1 \boldsymbol{v}_1 + \cdots + \alpha_k \boldsymbol{v}_k) + (\gamma_1 \boldsymbol{u}_1 + \cdots + \gamma_1 \boldsymbol{u}_l),$$

where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_k)$ belongs to the k-dimensional simplex

$$\mathcal{S}_k = \left\{ \boldsymbol{\alpha}; 0 \leq \alpha_j \leq 1 \, \forall \, j; \sum_{j=1}^{k} \alpha_j = 1 \right\}$$

and $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_l)'$ is a vector of positive coefficients. Suppose $\boldsymbol{V}$ and $\boldsymbol{U}$ are $p \times k$ and $p \times l$ matrices whose columns are the vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ and $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_l$, respectively. Then the MW decomposition of any vector $\boldsymbol{\theta}$ in $\wp$ will be given as

$$\boldsymbol{\theta} = \boldsymbol{V}\boldsymbol{\alpha} + \boldsymbol{U}\boldsymbol{\gamma}. \tag{3}$$

The vector $\boldsymbol{\theta}$ denotes any general parameter, but in our example, it stands for the mean parameter, $\boldsymbol{\mu}$. Returning to our illustration, let $p = 2$ and consider prior knowledge that restricts the mean, $\boldsymbol{\mu}$, with the constraint space defined in (1), where

$$\boldsymbol{A} = \begin{bmatrix} 1 & -1 \\ 0.1 & -1 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 90 \\ 1 \end{bmatrix}. \tag{4}$$

Applying the MW decomposition (for details regarding computation, see Neuman n.d.) on $\mathscr{M}$, we find $\boldsymbol{V}$ and $\boldsymbol{U}$ in (3),

$$\boldsymbol{V} = \begin{bmatrix} 0 & 10 & 98.89 \\ 0 & 0 & 8.89 \end{bmatrix}, \quad \boldsymbol{U} = \begin{bmatrix} 0 & 0.5 \\ 1 & 0.5 \end{bmatrix}.$$

Therefore, the parameter $\boldsymbol{\mu}$ may be represented as $\boldsymbol{\mu}(\boldsymbol{\alpha}, \boldsymbol{\gamma}) = \boldsymbol{V\alpha} + \boldsymbol{U\gamma}$ according to (3). Figure 2 provides geometric insight for the matrices $\boldsymbol{V}$ and $\boldsymbol{U}$ whose columns are the extreme points and the extreme directions, respectively. Every point in the constrained space, the shaded area, can be represented by a convex combination of the extreme points and a linear combination, with positive coefficients, of the extreme directions. Starting from a point in the simplex, the dark shaded area, the target point in the polyhedral, marked by the star, can be obtained as the sum of the two extreme vectors, properly scaled. The choice of the initial point and the pair of vectors is not unique and hence the parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ are not identifiable. However, we are only interested in $\boldsymbol{\mu}$, which is identifiable. Of course the matrices $\boldsymbol{V}$ and $\boldsymbol{U}$ depend on $\boldsymbol{A}$ and $\boldsymbol{b}$. The reparameterization of $\boldsymbol{\mu}$ in terms of $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ is the key to invoking the prior on $\boldsymbol{\mu}$. The MW representation provides a reparameterization of $\boldsymbol{\mu}$ in terms of $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ that makes the parameters essentially unrestricted and thereby facilitates prior specification.

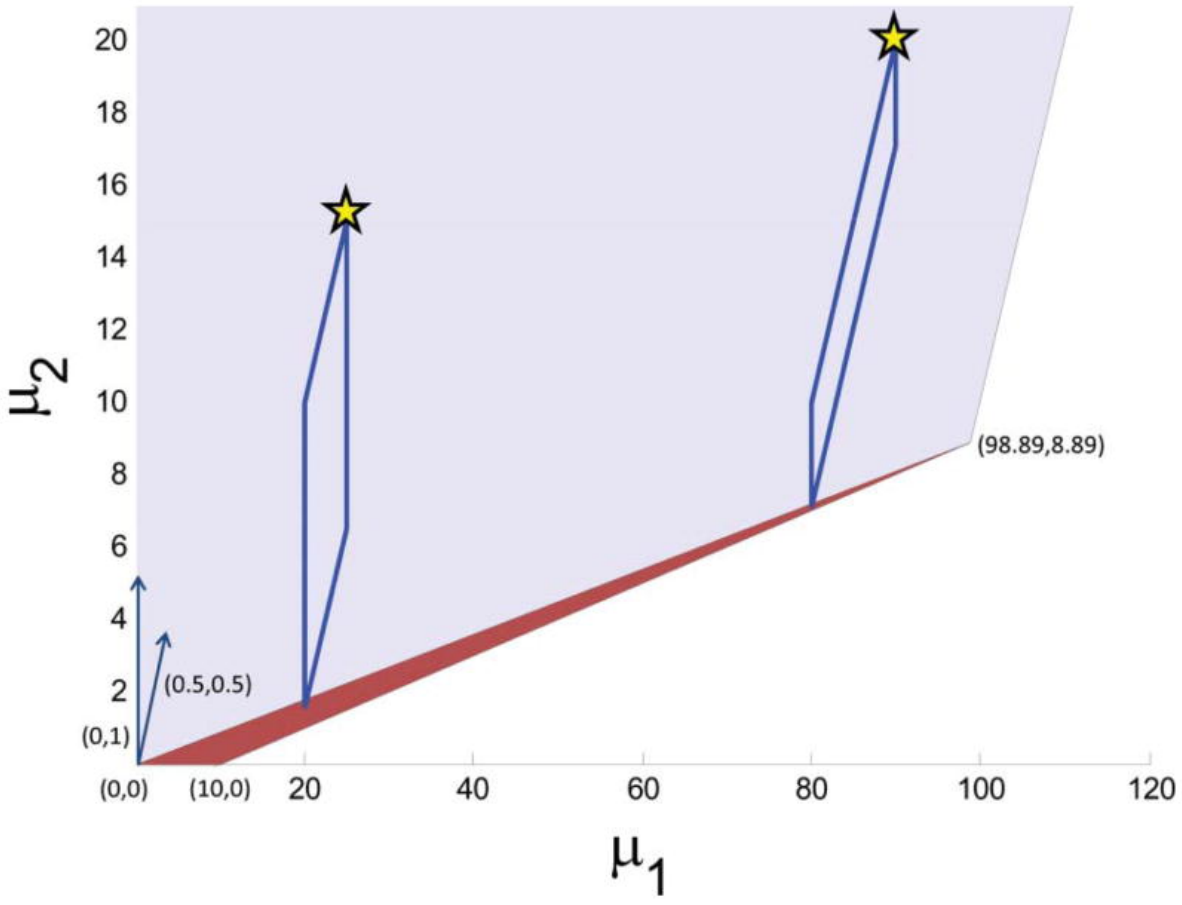


Figure 2

Illustration of Minkowski decomposition. The shaded area represents the constraint space, the dark shaded area represents the simplex, and the stars represent target points. The online version of this figure is in color.

2.1 Prior Specification

Once the matrices $\boldsymbol{V}$ and $\boldsymbol{U}$ are defined, we can invoke a prior distribution for $\boldsymbol{\mu}$ via prior distributions on $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ in (3). The flexibility of such prior specification becomes immediately clear because the parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ are essentially unconstrained. If needed, further transformation can be performed to make the range unrestricted, but given the abundance of known forms of densities that are fully supported on the simplex or on the positive quadrant, the need for a further transformation seems minimal. We will investigate the performance of priors that are probably the most obvious choices in this set up. Suppose we use a prior density π_α on $\boldsymbol{\alpha}$ and π_γ on $\boldsymbol{\gamma}$. Then the MW prior induced on $\boldsymbol{\mu}$ will be denoted as $p^{(\mathrm{MW})}_{\boldsymbol{V},\boldsymbol{U}}(\pi_\alpha, \pi_\gamma)$ or simply as $p^{(\mathrm{MW})}(\pi_\alpha, \pi_\gamma)$ if the matrices $\boldsymbol{V}$ and $\boldsymbol{U}$ are clear from the context. With slight abuse of notation, we will let $p^{(\mathrm{MW})}(\pi_\alpha, \pi_\gamma)$ also denote the product prior on $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ given by π_α and π_γ.

While a number of different prior distributions can be used, we choose the Dirichlet distribution for $\boldsymbol{\alpha}$ because it has support on the simplex and independent gamma distributions for elements of $\boldsymbol{\gamma}$ because they have support on the positive quadrant. The choice of the hyperparameters will depend on prior knowledge of the researchers. Weakly informative priors on the constraint space can be developed if a researcher has limited prior knowledge beyond the domain of restrictions for the parameters. In the context of the illustrative example, the choice of the Dirichlet parameters will determine the probable starting locations for the two vectors being summed, and the choice of the gamma parameters will determine the length of the vectors (see Figure 2). Therefore, weakly informative priors will allow for the starting location to be uniformly distributed on each dimension of the parameter space and for the length of the vectors to have a large variance.

We assume that $\boldsymbol{\alpha}$, $\boldsymbol{\gamma}$, and $\boldsymbol{\Sigma}$ are a priori mutually independent. Therefore, the joint prior density of $\boldsymbol{\alpha}$, $\boldsymbol{\gamma}$, and $\boldsymbol{\Sigma}$ is

$$p(\boldsymbol{\alpha}, \boldsymbol{\gamma}, \sum) = p^{(\mathrm{MW})}(\pi_\alpha, \pi_\gamma)p(\sum),$$

where π_α is the Dirichlet density and π_γ is the product of gamma densities. Since $\boldsymbol{\Sigma}$ is a priori independent of $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$, we choose an inverse Wishart prior for $\boldsymbol{\Sigma}$ for conjugacy. Since the inverse Wishart prior is known to be sensitive to the choice of the degrees of freedom (df), we also explored the use of the normal prior proposed by Leonard and Hsu (1992). The results closely match that with the inverse Wishart prior distributions, but we observed slower convergence of some of the MCMC chains in implementing the normal prior. Therefore, under this simple simulation scenario, we chose the inverse Wishart prior due to its nice computational properties. Thus we have the following prior distributions,

$$\begin{aligned} \textstyle\sum &\sim \mathrm{IW}(\boldsymbol{\Psi}, \kappa), \\ \boldsymbol{\alpha} &\sim \mathrm{Dir}(M, a_1, \ldots, a_k), \\ \boldsymbol{\gamma} &\sim \textstyle\prod_{i=1}^{l} \mathcal{G}(g_{1i}, g_{2i}), \end{aligned} \tag{5}$$

where $\mathrm{IW}(\boldsymbol{\Psi}, \kappa)$ is the inverse Wishart density with scale matrix $\boldsymbol{\Psi}$ and degrees of freedom κ, $\mathrm{Dir}(M, a_1, \ldots, a_k)$ is the density of the $(k-1)$ dimensional Dirichlet distribution with parameters $M > 0$ and $(a_1, \ldots, a_k) \in \mathcal{S}_{k-1}$, and $\mathcal{G}(g_1, g_2)$ denotes the Gamma density with parameters $g_1 > 0$ and $g_2 > 0$.

The choice of hyperparameters will be an important factor in determining the influence the prior will exert on the final analysis. The geometry of the constraint set and the MW representation provide some guidelines toward the choice of these parameters. The shape of the constraint space may automatically induce strong correlation among the coordinates of $\boldsymbol{\mu}$. However, given the particular shape, it will be prudent to have some strategy toward choosing the prior in a principled manner. As mentioned before, selection of a particular point, $\boldsymbol{\mu}$, is equivalent to choosing a starting point within the convex hull of the extreme points and then selecting the different scales on the set of extreme directions. To choose the initial point in a somewhat uniform manner within the convex hull, one may experiment with the

Dirichlet parameter once $\boldsymbol{V}$ has been determined. The weights for the direction vectors can be chosen to have large variance to achieve a somewhat diffused prior. Due to the constraints, having a diffused prior along each coordinate of $\boldsymbol{\gamma}$ does not necessarily produce a diffused prior over the restricted set. To have sufficient mass near the boundary, that is to move more toward one direction from the starting position than the other direction, negatively correlated coordinates may be considered. In the Simulation section, we evaluate the sensitivity of the procedure with respect to hyperparameter specification through a very limited study in the context of the illustrative example. More extensive simulations may provide more insight to the interplay between the prior parameters and the constraints and give us a better understanding of the mechanisms of prior specification for the MW priors.

3. POSTERIOR COMPUTATION AND SIMULATION

Using the likelihood defined in (2) and the prior distributions defined in (5), we are able to generate samples from the posterior distributions using the Metropolis–Hastings and Gibbs sampler formulation of Gelfand, Smith, and Lee (1992). To evaluate the performance of our approach, we analyze data in a limited simulation study under the scenario of the illustrative example. In particular, we consider the case where the parameters are as follows,

$$\boldsymbol{\mu} = \begin{bmatrix} 100.1 \\ 10.5 \end{bmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} 5 & 0.2 \\ 0.2 & 5 \end{bmatrix}.$$

The parameter value for $\boldsymbol{\mu}$ satisfies the constraints given in (1) with $\boldsymbol{A}$ and $\boldsymbol{b}$ specified in (4). The sample size is chosen to be similar to the sample size in the BioCycle data, $n = 250$. We generate the outcomes, $\boldsymbol{y}_i \overset{\text{iid}}{\sim} N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ for $i = 1, 2, \ldots, n$. Furthermore, we choose $\kappa = 10$ and $\boldsymbol{\Psi} = \kappa \times \boldsymbol{I}_2$, where $\boldsymbol{I}_p$ is the p dimensional identity matrix.

3.1 Simulation

We investigate the accuracy and sensitivity of the estimation procedure for several choices of the Dirichlet and gamma parameters. To choose the Dirichlet parameters, we parameterize the distribution as Dir (M, a_1, a_2, a_3), where M is chosen from the set $\{0.01, 0.1, 1, 2\}$ and (a_1, a_2, a_3) is chosen in a representative manner from the simplex. Specifically, (a_1, a_2, a_3) is chosen as an element of the set $\{(0.45, 0.05, 0.5), (0.05, 0.05, 0.9), (0.05, 0.9, 0.05), (0.9, 0.05, 0.05), (0.33, 0.33, 0.34)\}$. For the Dirichlet distribution, larger choices of M will place more prior mass near the mean of the Dirichlet distribution, and the choice of a_1, a_2, and a_3 will weight the density on each vertex of the simplex. We choose the gamma priors to be independent $\mathcal{G}(g_1, g_2)$ with mean $g_1 g_2$ and variance $g_1 g_2^2$, and (g_1, g_2) is chosen as an element of the set $\{(1,4),(1,6),(1,8)\}$. Posterior samples are generated using the Metropolis–Hastings within Gibbs loop algorithm. To generate posterior estimates, we generate a chain of size 5000 following a burn-in of 5000. The number of Monte Carlo replications for each scenario is 1000. Figure 3 shows the root mean squared error (RMSE) for both coordinates of the mean. We observe that the estimation procedure is robust to different choices of the prior distribution for $\boldsymbol{\alpha}$ (the

RMSE is between 0.12 and 0.16 for all simulated priors). Additionally, since the true mean is near a simplex vertex, the estimation procedure does better with choices of M that are small and thus place more prior density near the vertices.

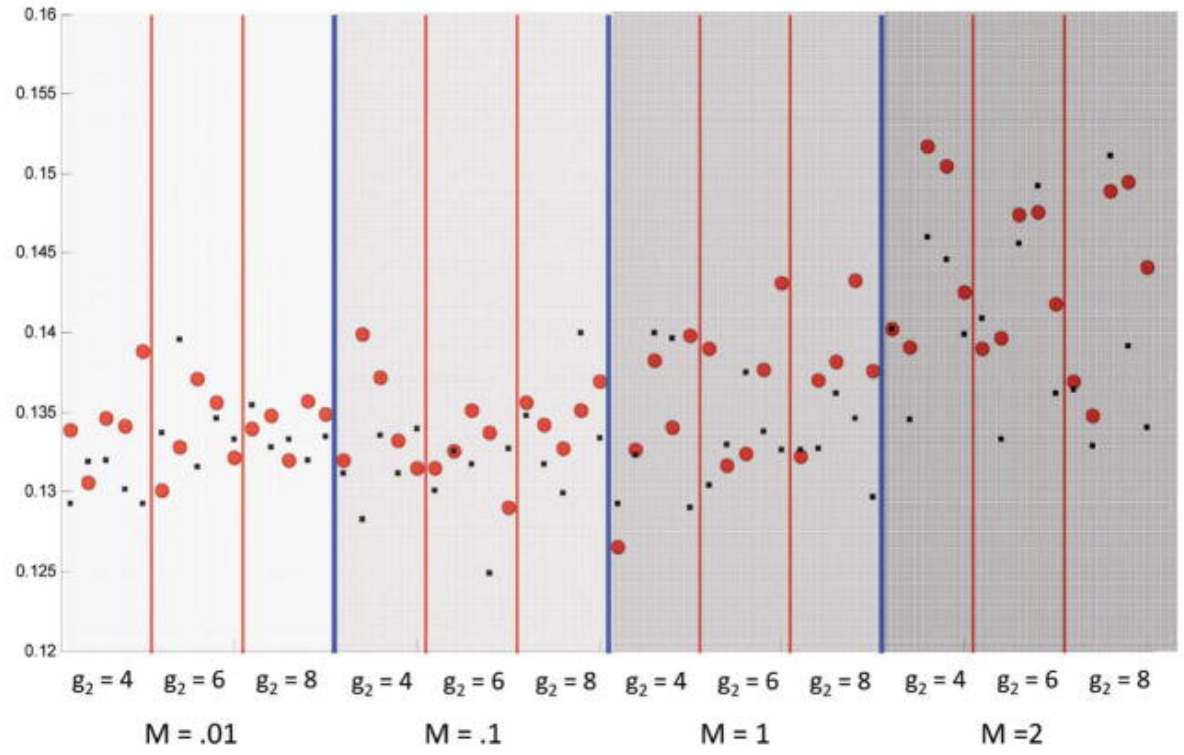


Figure 3

RMSE of posterior estimates using different prior distributions. M, a_1, a_2, a_3 denote Dirichlet parameters and g_1, g_2 denote gamma parameters. For all simulations $g_1 = 1$. There are 12 sections of the figure identified by the value of M and g_2. Within each partition, each point represents different choices of (a_1, a_2, a_3) in the following order (0.45, 0.05, 0.5), (0.05, 0.05, 0.9), (0.05, 0.9, 0.05), (0.9, 0.05, 0.05), (0.33, 0.33, 0.34). The circle dots correspond to the RSME for μ_1, and the small squares for μ_2. The online version of this figure is in color.

3.2 Comparison With Existing Methods

As an alternative to our estimation procedure, one could use an unrestricted prior and discard inconsistent draws from the posterior distribution, as proposed by Gelfand, Smith, and Lee (1992). To explore this option, we used the following unrestricted prior for $\boldsymbol{\mu}$,

$$\boldsymbol{\mu} \sim N\left(\begin{bmatrix} 100 \\ 10 \end{bmatrix}, \begin{bmatrix} 100 & 0 \\ 0 & 10 \end{bmatrix}\right).$$

We simulated 1000 different samples and found posterior distributions using Gibbs chains of length 5000 with a burn-in of 5000. Of the 1000 chains, 16 chains needed more than 20,000 posterior draws to obtain 10,000 within the constraint space, 242 chains needed more than 11,000 draws, and 30 chains needed no extra draws. The median number of draws was 10,186. We note that the two-dimensional illustrative example is almost the simplest example with linear constraints. In more complicated situations, such as the real data analysis in the next section, high dimensional constraints will make this truncation approach very inefficient. Figure 4 is the contour plot of 1000 estimates of the mean.

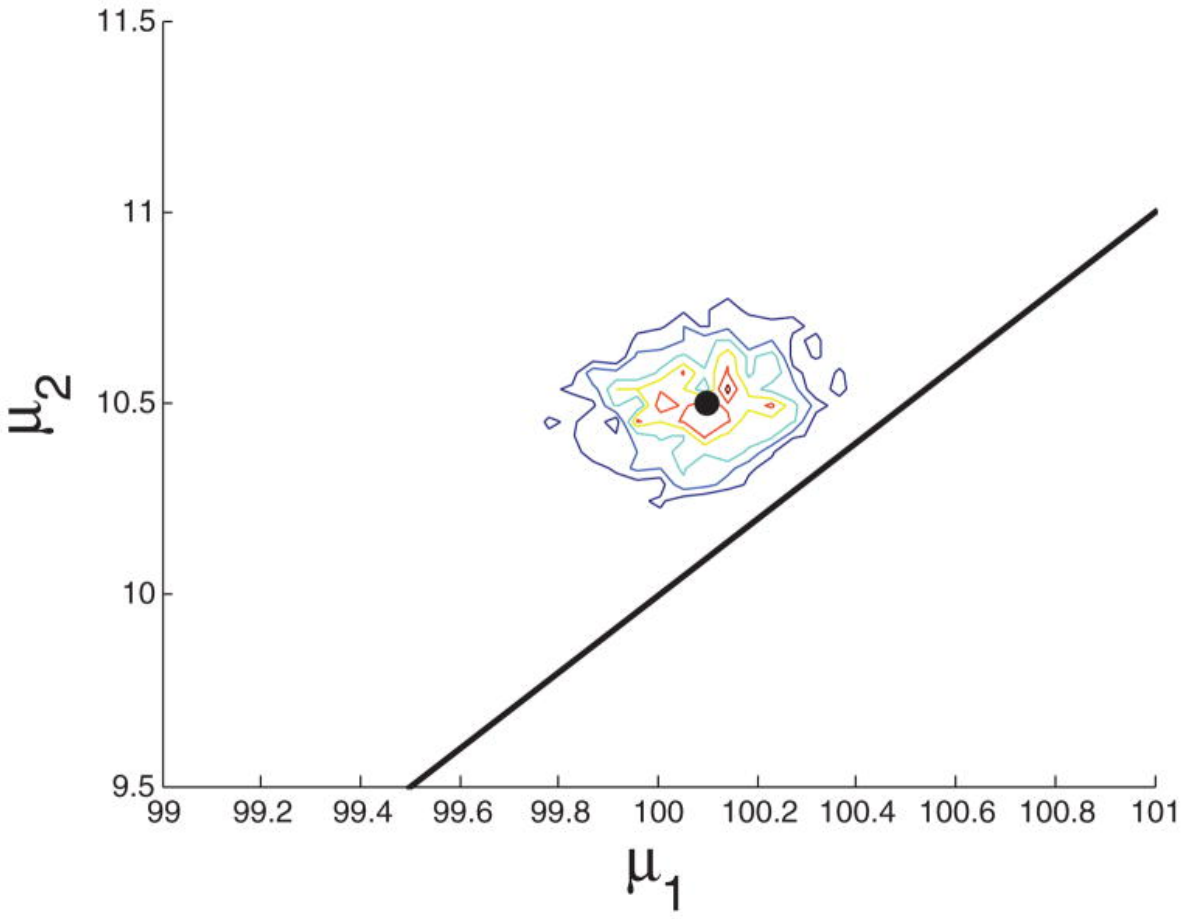
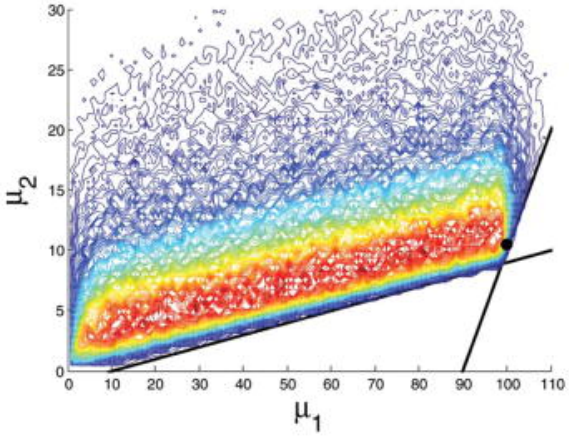


Figure 4

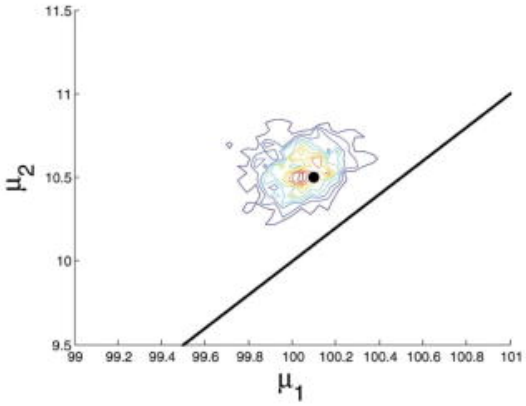
Contour plot of the estimated posterior means for $\boldsymbol{\mu}$ using a normal prior where posterior draws inconsistent with the constraint space are discarded. The black dot represents the location of the true mean used in the simulation. The online version of this figure is in color.

3.3 Guidelines

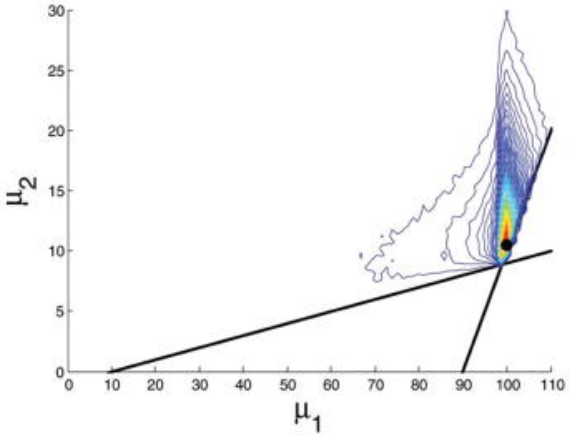
Recall from our discussion of Figure 2, the mean, $\boldsymbol{\mu}$, is parameterized as the sum of $\boldsymbol{V\alpha}$, which is the starting location, and $\boldsymbol{U\gamma}$, which is an appropriately scaled sum of vectors. One must carefully choose a prior distribution for $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ to obtain a prior distribution for the mean, $\boldsymbol{\mu}$, that is nearly uniform over the constraint space. The choice of the prior distribution of $\boldsymbol{\alpha}$ should provide nearly uniform starting locations on each dimension of the constraint space, and the choice of the prior distribution of $\boldsymbol{\gamma}$ should allow the length of the vectors to have a large variance. To obtain the parameters for the Dirichlet distribution that provide uniform starting locations on each dimension of the constraint space, we consider the area (or volume for more than two dimensions) around each simplex vertex. We select the Dirichlet parameters such that the distribution is more dense where the area is small and less dense where the area is large. Therefore we select the prior distribution, Dir (2, 0.45, 0.05, 0.5), for $\boldsymbol{\alpha}$. One can easily check that the selected prior distribution yields an approximately uniform distribution of starting locations on each dimension of the constraint space by taking a random sample from the prior distribution and plotting a histogram of $\boldsymbol{V\alpha}$ for each dimension. Selecting parameters for the gamma distribution that allow for the length of the vectors to have a large variance is straightforward. For each $\gamma_j, j = 1, 2$, we select the prior distribution, $\mathcal{G}$ (1, 4), which has mean 4 and variance 16. Since this example only has two dimensions, we are able to visualize the induced prior density of $\boldsymbol{\mu}(\boldsymbol{\alpha}, \boldsymbol{\gamma})$. In Figure 5(a), we provide the contour plot of 500,000 random samples from the prior distribution. This prior distribution for $\boldsymbol{\mu}$ is nearly evenly spread within the constraint space.



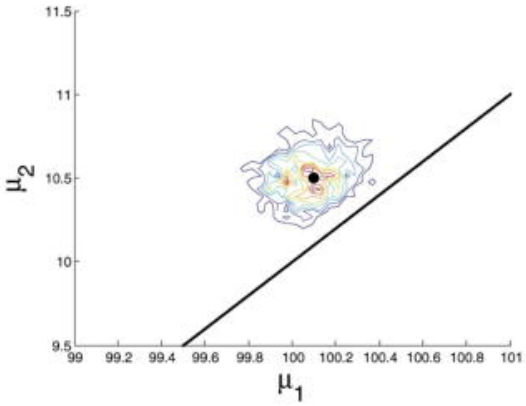
(a) $M, \boldsymbol{a} = (2, 0.45, 0.05, 0.5)$



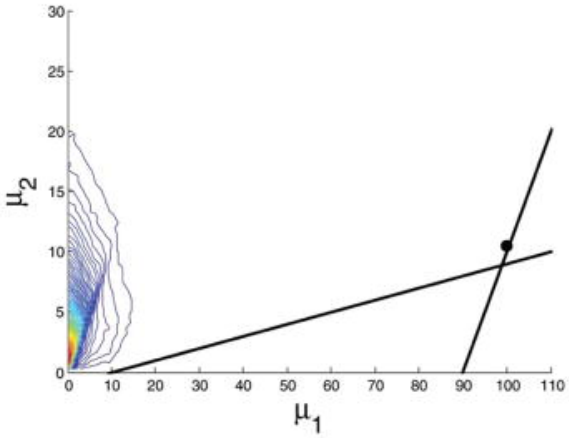
(b) $M, \boldsymbol{a} = (2, 0.45, 0.05, 0.5)$



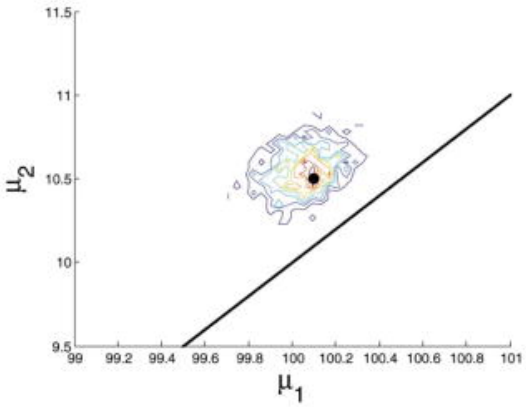
(c) $M, \boldsymbol{a} = (1, 0.05, 0.05, 0.9)$



(d) $M, \boldsymbol{a} = (1, 0.05, 0.05, 0.9)$



(e) $M, \boldsymbol{a} = (1, 0.9, 0.05, 0.05)$



(f) $M, \boldsymbol{a} = (1, 0.9, 0.05, 0.05)$

Open in a separate window

Figure 5

On the left are prior densities of the means, $\boldsymbol{\mu}$, when the Dirichlet distribution with parameters (M, $\boldsymbol{a}$) is the prior for $\boldsymbol{\alpha}$ and the gamma distribution with parameters (1, 4) is the prior for $\boldsymbol{\gamma}$. On the right are the corresponding samples from the posterior distribution. The black dot represents the location of the true mean used in the simulation. Notice that the figures on the left are scaled to focus in near the true mean. The online version of this figure is in color.

Using this prior distribution, Figure 5(b) plots the 1000 estimates of the mean from the posterior distribution. The true mean for this simulation study is close to the boundary of the constrained space. For the particular parameter combination in the figure, the estimates obtained from the MW approach do not cluster near the constraint boundary. This is because the chosen MW prior is concentrated inside the constraint space. This phenomenon results in a slight inward bias for the estimates; however, the RMSE is still quite small. If the researcher has prior knowledge about the location of the mean, an informative prior would put more mass toward the desired location.

To explore different prior distributions, we considered a very informative prior that put a lot of mass near the true mean parameter and a misspecified prior that put a very small amount of mass near the true mean parameter. Figure 5(c) shows the contour plot of random samples from the induced prior density of $\boldsymbol{\mu}(\boldsymbol{\alpha}, \boldsymbol{\gamma})$ when we select Dir (1, 0.05, 0.05, 0.9) for $\boldsymbol{\alpha}$ and independent $\mathcal{G}$ (1, 4) for each γ_j, j = 1, 2, and Figure 5(d) plots the 1000 posterior estimates of the mean. We see that when the mass of the prior is more concentrated in the location of the true mean, the posterior estimates cluster closer toward the boundary. Additionally, Figure 5(e) shows the induced prior density of $\boldsymbol{\mu}(\boldsymbol{\alpha}, \boldsymbol{\gamma})$ when we select Dir (1, 0.9, 0.05, 0.05) for $\boldsymbol{\alpha}$ and independent $\mathcal{G}$ (1, 4) for each $\gamma_j, j = 1, 2$. Though this is a very informative prior, it puts very little prior mass near the true mean parameter. Figure 5(f) plots the 1000 posterior estimates of the mean. These posterior estimates do not behave much worse than the less informative prior. Table 1 provides the mean, variance, and mode for each prior distribution of the mean, $\boldsymbol{\mu}$. The mean and variance are defined by, $E(\boldsymbol{\mu}) = \boldsymbol{V}E(\boldsymbol{\alpha}) + \boldsymbol{U}E(\boldsymbol{\gamma})$, $\text{var}(\boldsymbol{\mu}) = \boldsymbol{V}\text{var}(\boldsymbol{\alpha})\boldsymbol{V}' + \boldsymbol{U}\text{var}(\boldsymbol{\gamma})\boldsymbol{U}'$, and the mode is estimated from 500,000 random samples from the prior density of the mean, $\boldsymbol{\mu}$. Additionally, Table 1 tabulates the bias, RMSE, and the frequentist coverage of 90% credible intervals for both coordinates of the mean, $\boldsymbol{\mu}$, using the noninformative, informative, and misspecified priors.

Table 1

Comparing different hyperparameters. The true mean is $\boldsymbol{\mu} = [100.1, 10.5]'$. The prior for γ is $\mathcal{G}(1, 4)$, and the prior for $\boldsymbol{\alpha}$ is Dir(M, a_1, a_2, a_3). Coverage probability is for the 90% credible intervals

	Dirichlet parameter (M, a_1, a_2, a_3)		
	(2, 0.45, 0.05, 0.50) (μ_1, μ_2)	(1, 0.05, 0.05, 0.90) (μ_1, μ_2)	(1, 0.90, 0.05, 0.05) (μ_1, μ_2)
Prior mean	51.9444, 10.4444	91.500, 14.000	7.444, 6.444
Prior mode	73.126, 8.858	99.331, 9.867	0.333, 0.809
Prior variance	804.020, 26.584	401.931, 23.556	236.154, 21.877
Bias	−0.055, 0.049	−0.024, 0.018	−0.051, 0.050
RMSE	0.140, 0.140	0.133, 0.132	0.134, 0.140
Coverage probability	0.875, 0.876	0.896, 0.897	0.882, 0.876

Open in a separate window

4. ANALYSIS OF THE BIOCYCLE STUDY DATA

As discussed in the introduction, the BioCycle Study contains a prospective cohort of 259 premenopausal women who were followed for up to two menstrual cycles. Reproductive hormones were measured in serum samples at eight well-timed visits during each menstrual cycle. Figure 1 illustrates the expected hormonal curves and where the eight visits were anticipated to sample. Details of the BioCycle Study and its sampling scheme are published elsewhere (Wactawski-Wende et al. 2009). Using the data from the BioCycle Study, our interest is to investigate the relationship between markers of oxidative stress and reproductive hormones among normally menstruating women. Earlier studies have shown a possible relationship between reproductive hormones and oxidative stress, however the relationships have been inconclusive (Schisterman et al. 2010). Using biological knowledge about hormonal profiles, we build a model with the hormones as our response and F_2-isoprostanes, an oxygen-free radical known to be an agent for oxidative stress, as a covariate. Figure 6 illustrates the unadjusted relationship between the reproductive hormones and F_2-isoprostanes within the BioCycle Study. This figure suggests that higher levels of oxidative stress appear to be related to higher levels of estrogen and progesterone and lower levels of LH and FSH. Using a constraint space involving mean log hormone levels, we hope to build an efficient model to detect these relationships if they exist.

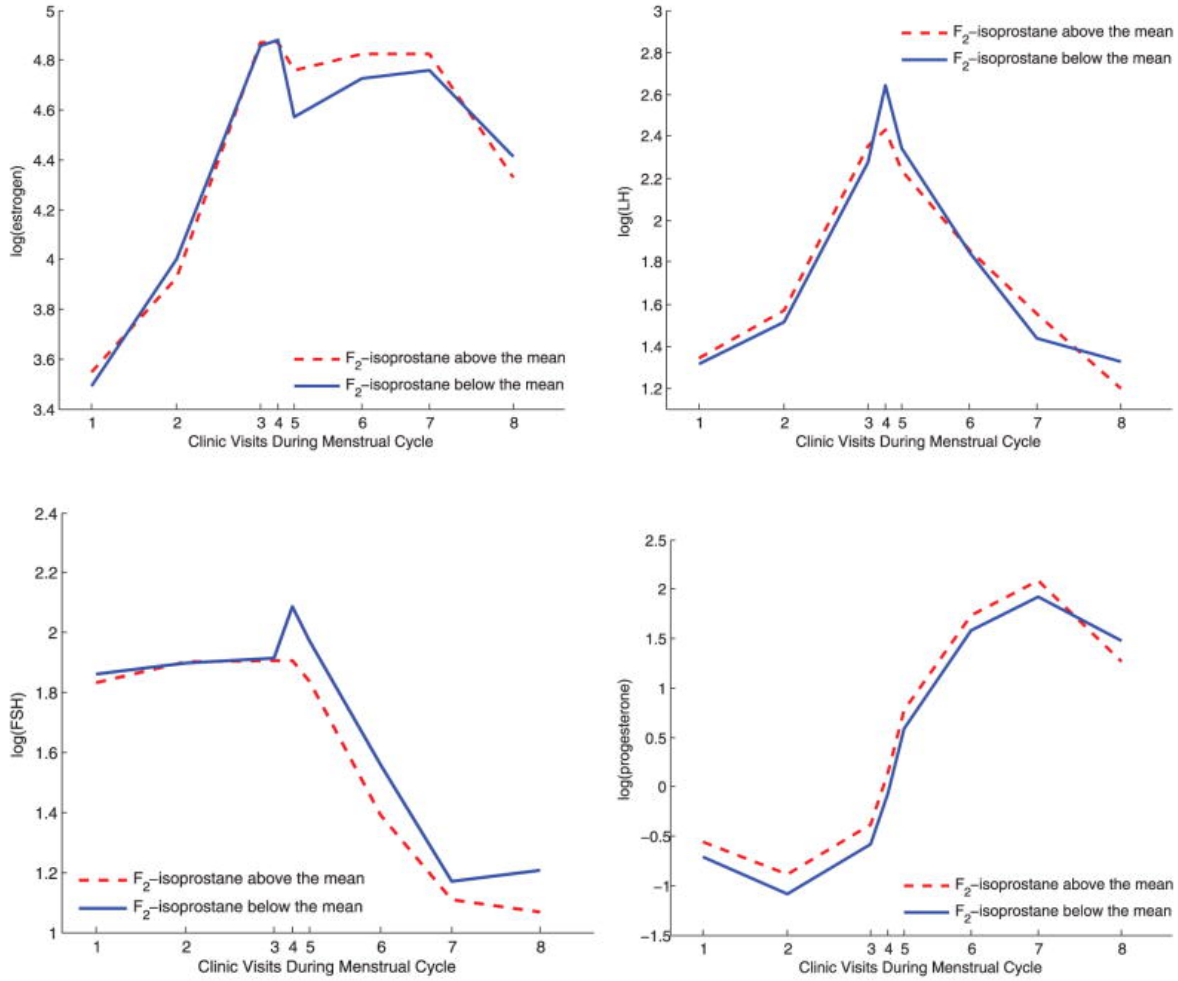


Figure 6

Unadjusted plots of the log observed hormone levels by level of F_2-isoprostanes (above and below the mean, mean = 48.61 pg/mL). The online version of this figure is in color.

Roy et al. (2011) used the BioCycle Study to explore the relationship between reproductive hormones and oxidative stress in a multivariate linear mixed model with constrained mean parameters. To enforce constraints on their parameters, they used a projection approach where the posterior draws obtained using an unconstrained prior were projected back to the constrained space within the Gibbs sampling iteration (Gunn and Dunson 2005). We apply a similar multivariate linear mixed model, however we constrain our parameters with the MW prior. We define y_{iht} to be the observation on the ith woman, $i = 1, \ldots, n$, of the hth hormone, $h \in \{1, \ldots, H\}$, at visit t, $t \in \{1, \ldots, T\}$. In the BioCycle Study, we are interested in characterizing four hormones ($H = 4$); therefore to ease notational burden, we define $h \in \{E, L, F, P\}$, which represents estrogen, LH, FSH, and progesterone, respectively. Our model is given by

$$y_{iht} = \mu_{ht} + \boldsymbol{w}'_{it}\boldsymbol{\phi}_h + \xi_{ih} + \zeta_{it} + \varepsilon_{iht}, \tag{6}$$

where μ_{ht} is the mean hormone level of the population for hormone h at visit t, $\boldsymbol{w}_{it}$ is a $q \times 1$ vector of covariates (including both baseline and time-varying covariates) for subject i at visit t, $\boldsymbol{\phi}_h$ is a $q \times 1$ vector of covariate fixed effects, and ξ_{ih} and ζ_{it} are the hormone-specific and visit-specific subject random effects, respectively. We use these random effects to account for correlated observations arising from multiple hormone measurements per woman at the same time point and multiple visits across the cycle for each hormone. The key covariate of interest is F_2-isoprostanes (pg/mL). To adjust for possible confounding effects, three baseline covariates and four time-varying covariates are included in the model. The baseline covariates are age (in years), race (Caucasian or other indicated by 0 and 1, respectively), and age at menarche (in years), and the time-varying covariates are γ-vitamin E (μg/mL), beta-carotene (μg/mL), total cholesterol (mg/dL), and homocysteine levels (μmol/L) (Schisterman et al. 2010). Additionally, in keeping with Schisterman et al. (2010), we log transform beta-carotene, γ-vitamin E, and F_2-isoprostanes, and to assist in convergence, we standardize all covariates. Although data were collected on 250 participants for two cycles (9 women had data collected for only one cycle), we restrict the analysis to the second cycle because the second cycle has fewer missing values and is more consistent with the known hormonal behavior during the different phases: menses, follicular phase, ovulation, and luteal phases. We exclude 18 women from the analysis because they were classified as anovulatory for the second observed cycle (Gaskins et al. 2009). Also, we exclude 2 women for missing age at menarche (in years), 5 women for missing beta-carotene and γ-vitamin E for all timed visits in their second cycle, and 3 women for missing beta-carotene for all timed visits in their second cycle. After exclusions, for our analysis $n = 222$, $q = 8$, $H = 4$, $T = 8$. However, not all women had all T scheduled visits, and hence one must develop some notation, tedious at times, to write the model in a more compact form.

Let n_i denote the number of visits (repeated measures) for the ith subject, and $\boldsymbol{t}_i = (t_{i(1)}, \ldots, t_{i(n_i)})$ denote the observed visit times with $t_{i(j)} \in \{1, \ldots, T\}$. Let $\boldsymbol{I}_a$ denote the $a \times a$ identity matrix, $\boldsymbol{i}_{j,T}$ denote the jth row of $\boldsymbol{I}_T$, and $\mathbf{1}_b$ denote the $b \times 1$ vector of 1s. In addition, let $\boldsymbol{I}(\boldsymbol{t}_i)$ denote the $n_i \times T$ matrix given by $(\boldsymbol{i}'_{t_{i(1)},T} : \cdots : \boldsymbol{i}'_{t_{i(n_i)},T})'$. Then the model for the observations for the ith subject can be written as

$$\boldsymbol{y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i^{\mathcal{H}}\boldsymbol{\xi}_i + \boldsymbol{Z}_i^{\mathcal{T}}\boldsymbol{\zeta}_i + \boldsymbol{\varepsilon}_i, \tag{7}$$

where $\boldsymbol{y}_i = (y_{i\,Et_{i(1)}}, \ldots, y_{i\,Eti(n_i)}, \ldots, y_{i\,Pt_{i_{(1)}}}, \ldots, y_{i\,Pt_i(n^i)})'$ is the vector of observations for woman i, $\boldsymbol{\xi}_i = (\xi_{iE}, \xi_{iL}, \xi_{iF}, \xi_{iP})$ and $\boldsymbol{\zeta}_i = (\zeta_{i1}, \zeta_{i2}, \ldots, \zeta_{iT})$ are the vectors of hormone and time specific subject random effects for woman i, respectively. In addition, $\boldsymbol{X}_i = (\boldsymbol{M}_i, \boldsymbol{W}_i)$ is a $n_i\,H \times (TH + qH)$ matrix where $\boldsymbol{M}_i = (\boldsymbol{I}_H \otimes \boldsymbol{I}(\boldsymbol{t}_i))$, $\boldsymbol{W}_i = (\boldsymbol{I}_H \otimes \boldsymbol{w}_i)$, $\boldsymbol{w}_i$ is a $n_i \times q$ matrix of the q covariate values for the ith women during the n_i visits, and the symbol $\otimes$ is the Kronecker product. Furthermore, $\boldsymbol{Z}_i^{\mathcal{H}} = (\boldsymbol{I}_H \otimes \mathbf{1}(\boldsymbol{t}_i))$ and $\boldsymbol{Z}_i^{\mathcal{T}} = (\mathbf{1}_H \otimes \boldsymbol{I}(\boldsymbol{t}_i))$. We will make standard distributional assumption on the random effects,

$$\boldsymbol{\xi}_i^{\mathcal{H}} \sim N(\mathbf{0}, \boldsymbol{D}^{\mathcal{H}}), \qquad \boldsymbol{\zeta}_i^{\mathcal{T}} \sim N(\mathbf{0}, \boldsymbol{D}^{\mathcal{T}}),$$

and the random effects are assumed to be independent across subjects. The errors are assumed to be independently and identically distributed as $N(0, \sigma_e^2)$) and independent of the random effects. The fixed effect parameter $\boldsymbol{\beta}$ comprises the TH mean population hormone levels $\boldsymbol{\mu} = (\mu_{E,1}, \ldots, \mu_{E,T}, \ldots, \mu_{P,1}, \ldots, \mu_{P,T})'$ and the qH covariate effects $\boldsymbol{\phi} = (\boldsymbol{\phi}_E, \ldots, \boldsymbol{\phi}_P)'$. Thus, $\boldsymbol{\beta}$ is a $(TH + qH) \times 1$ vector of population parameters.

4.1 Constraints for the BioCycle Data

Specification of the constraints for the BioCycle Study is based on our knowledge of the timing of events across the menstrual cycle and the interdependencies between the reproductive hormones estrogen, FSH, LH, and progesterone. We use the general depiction of the hormonal patterns given by Figure 1, the discussion thereafter, and the expert opinion of reproductive epidemiologists to formulate the linear inequality constraints.

We formulated six linear constraints involving nine mean parameters. First, we know that the LH must surge to trigger ovulation, as seen in Figure 1 (Gilbert 2006). The surge can be defined when the LH level reaches at least 1.5 times the mean preovulatory level observed during the follicular phase (Scolaro, Lloyd, and Helms 2008). Therefore we expect that at the fourth visit, $t = 4$, the measured LH is at least 1.5 times the measured LH at the second visit, $t = 2$. Also, we know that LH must increase and eventually decrease to the pre-surge level. However the return to the pre-surge level may happen after the final visit and only a fraction of the total decrease may occur within the range of the visit days. Therefore we expect that the measured LH at the surge should be at least 1.1 times (10% more than) the values at the last visit, $t = 8$. Since we log transform the hormone measurements, these qualitative constraints translate into linear constraints. Additionally, we know that during the follicular, phase estrogen suppresses production of LH until a threshold level is reached, at which point LH production is stimulated and then peaks (Hall 2009). Therefore the estrogen measured during menses, $t = 1$, must be less than the estrogen measured near the LH peak, $t = 4$. Additionally, we know that progesterone levels rise after ovulation as the corpus luteum forms and secretes progesterone (Hall 2009). Therefore, we expect that the measured progesterone level at $t = 7$ is more than the progesterone measured at $t = 1$ and $t = 2$. Finally, if the egg is not fertilized, then the corpus luteum begins to degenerate and hormone secretion decreases causing progesterone concentrations to fall and initiate a new cycle (Hall 2009). Therefore the measured progesterone at $t = 8$ is expected to be less than at $t = 7$.

Using these known biological hormonal patterns, we develop constraints for the relevant hormone means, $(\mu_{E1}, \mu_{E4}, \mu_{L2}, \mu_{L4}, \mu_{L8}, \mu_{P1}, \mu_{P2}, \mu_{P7}, \mu_{P8})$. We partition $\boldsymbol{\beta}$ into constrained and unconstrained parameters, $\boldsymbol{\beta} = [\boldsymbol{\beta}^C \boldsymbol{\beta}^U]'$. Furthermore, we partition $\boldsymbol{\beta}^C$ into independent constraints for each hormone, $\boldsymbol{\beta}^C = [\boldsymbol{\beta}^{C_E} \boldsymbol{\beta}^{C_L} \boldsymbol{\beta}^{C_P}]$. Notice this partition is helpful because the complexity of the MW decomposition will grow only as fast as that of the constraint with the largest number of extreme points (k) in the partitioned form. Since we log transform the hormone measurements, the means may be positive or negative, which increases the computational difficultly of the MW decomposition. Therefore, to ease computation, we transform $\boldsymbol{\beta}^C$ to the positive space by defining a maximum for the constrained means,

$\boldsymbol{\beta}^C < \boldsymbol{M}_\beta$, where $\boldsymbol{M}_\beta = 10 \times \mathbf{1}_9$. In many biological applications, it is reasonable to define a maximum for parameters. Once we define the maximum, we define $\boldsymbol{\beta}^{C*} > \mathbf{0}$, such that $\boldsymbol{\beta}^C = \boldsymbol{M}_\beta - \boldsymbol{\beta}^{C*}$. The following is the constraint space for $\boldsymbol{\beta}^{C*}$,

$$
\begin{aligned}
\mathscr{M} &= \{\boldsymbol{\beta}^{C*} = [\boldsymbol{\beta}^{C_E*}\boldsymbol{\beta}^{C_L*}\boldsymbol{\beta}^{C_P*}]' : \boldsymbol{A}^h\boldsymbol{\beta}^{C_h*} < \boldsymbol{b}^h, \boldsymbol{\beta}^{C_h*} > \mathbf{0}, h \in \{E, L, P\}\}, \\
\boldsymbol{A}^E &= \begin{bmatrix} -1 & 1 \end{bmatrix}, \quad \boldsymbol{b}^E = [0], \\
\boldsymbol{A}^L &= \begin{bmatrix} -1 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix}, \quad \boldsymbol{b}^L = \begin{bmatrix} \log(1.5) \\ \log(1.1) \end{bmatrix}, \\
\boldsymbol{A}^P &= \begin{bmatrix} -1 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}, \quad \boldsymbol{b}^P = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.
\end{aligned}
$$

Applying the Minkowski decomposition to each partition of $\mathscr{M}$, we find the corresponding $\boldsymbol{V}^h$ and $\boldsymbol{U}^h$ matrices from (3),

$$
\begin{aligned}
\boldsymbol{V}^E &= \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{U}^E = \begin{bmatrix} 1 & 0.5 \\ 0 & 0.5 \end{bmatrix}. \\
\boldsymbol{V}^L &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0.095 & 0.406 \\ 0 & 0 & 0.310 \end{bmatrix}, \quad \boldsymbol{U}^L = \begin{bmatrix} 1 & 0.333 & 1 \\ 0 & 0.333 & 0 \\ 1 & 0.333 & 0 \end{bmatrix}. \\
\boldsymbol{V}^P &= \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{U}^P = \begin{bmatrix} 0 & 0 & 0.25 & 1 \\ 0 & 1 & 0.25 & 0 \\ 0 & 0 & 0.25 & 0 \\ 1 & 0 & 0.25 & 0 \end{bmatrix}.
\end{aligned}
$$

4.2 MW Prior and Posterior Inference

Once the constraints are well defined using prior knowledge regarding hormone dynamics, we complete our model by defining the prior distributions of the model parameters. We use a novel reparameterization proposed by Chen and Dunson (2003) for the covariance matrices of the random effects, $\boldsymbol{D}^{\mathscr{H}}$ and $\boldsymbol{D}^{\mathscr{T}}$, which accounts for the uncertainty in specifying the random effect correlation structure. More specifically, the covariance matrices of the random effects are parameterized by the Cholesky decomposition, $\boldsymbol{D}^{\mathscr{H}} = \boldsymbol{L}^{\mathscr{H}}\boldsymbol{L}^{\mathscr{H}\prime}$ and $\boldsymbol{D}^{\mathscr{T}} = \boldsymbol{L}^{\mathscr{T}}\boldsymbol{L}^{\mathscr{T}\prime}$, where $\boldsymbol{L}^{\mathscr{H}}$ and $\boldsymbol{L}^{\mathscr{T}}$ are lower triangular matrices. The random effects are thus written as $\boldsymbol{\xi}_i = \boldsymbol{L}^{\mathscr{H}}\boldsymbol{\delta}_i^{\mathscr{H}}$ and $\boldsymbol{\zeta}_i = \boldsymbol{L}^{\mathscr{T}}\boldsymbol{\delta}_i^{\mathscr{T}}$, with both $\boldsymbol{\delta}_i^{\mathscr{H}}$ and $\boldsymbol{\delta}_i^{\mathscr{T}}$ comprising independent standard normal variates. The Cholesky square roots are further decomposed

as $\boldsymbol{L}^{\mathcal{H}} = \boldsymbol{\Lambda}^{\mathcal{H}}\,\boldsymbol{\Gamma}^{\mathcal{H}}$ and $\boldsymbol{L}^{\mathcal{T}} = \boldsymbol{\Lambda}^{\mathcal{T}}\,\boldsymbol{\Gamma}^{\mathcal{T}}$, where the $\boldsymbol{\Lambda}$'s are diagonal matrices with nonnegative diagonal entries and $\boldsymbol{\Gamma}$'s are lower triangular matrices with ones along the diagonal, with the restriction that if the ith diagonal entry of $\boldsymbol{\Lambda}$ is zero, then all entries in the ith column and in the ith row of $\boldsymbol{\Gamma}$ are zero. Consequently, model (7) can be rewritten as

$$\boldsymbol{y}_i = \boldsymbol{X}_i^C\boldsymbol{\beta}^C + \boldsymbol{X}_i^U\boldsymbol{\beta}^U + \boldsymbol{Z}_i^{\mathcal{H}}\boldsymbol{\Lambda}^{\mathcal{H}}\boldsymbol{\Gamma}^{\mathcal{H}}\boldsymbol{\delta}^{\mathcal{H}} + \boldsymbol{Z}_i^{\mathcal{T}}\boldsymbol{\Lambda}^{\mathcal{T}}\boldsymbol{\Gamma}^{\mathcal{T}}\boldsymbol{\delta}^{\mathcal{T}} + \boldsymbol{\varepsilon}_i. \tag{8}$$

The parameters corresponding to hormone-specific subject random effects are then $\boldsymbol{\lambda}^{\mathcal{H}} = (\lambda_E^{\mathcal{H}}, \lambda_L^{\mathcal{H}}, \lambda_F^{\mathcal{H}}, \lambda_P^{\mathcal{H}})'$, where $\lambda_i^{\mathcal{H}}$ is the ith diagonal entry of $\boldsymbol{\Lambda}^{\mathcal{H}}$, and $\boldsymbol{\gamma}^{\mathcal{H}} = \{\gamma_{ml}^{\mathcal{H}}, m = 2, \ldots, H; l = 1, \ldots, m-1\}'$, where $\gamma_{ml}^{\mathcal{H}}$ is the (m, l)th entry of $\boldsymbol{\Gamma}^{\mathcal{H}}$. Similarly, we define $\boldsymbol{\lambda}^{\mathcal{T}}$ and $\boldsymbol{\gamma}^{\mathcal{T}}$ corresponding to the visit specific subject random effects. We denote the entire set of parameters as $\boldsymbol{\theta} = (\boldsymbol{\beta}^{C\prime}, \boldsymbol{\beta}^{U\prime}, \boldsymbol{\lambda}^{\mathcal{H}\prime}, \boldsymbol{\gamma}^{\mathcal{H}\prime}, \boldsymbol{\lambda}^{\mathcal{T}\prime}, \boldsymbol{\gamma}^{\mathcal{T}\prime}, \sigma_e^2)'$.

To complete the prior specifications, we make the assumption that the prior distributions for the fixed parameters $\boldsymbol{\beta}^C$ and $\boldsymbol{\beta}^U$, the error precision σ_e^2, and the variance components $\boldsymbol{D}^{\mathcal{H}}$ and $\boldsymbol{D}^{\mathcal{T}}$ are mutually independent:

$$p(\boldsymbol{\theta}) = p(\boldsymbol{\beta}^C)p(\boldsymbol{\beta}^U)p\left(\sigma_e^2\right)p(\boldsymbol{\lambda}^{\mathcal{H}}, \boldsymbol{\gamma}^{\mathcal{H}})p(\boldsymbol{\lambda}^{\mathcal{T}}, \boldsymbol{\gamma}^{\mathcal{T}}).$$

Following standard convention, we choose conjugate priors when possible to ease posterior sampling. Following these principles, we propose the following priors. For the unconstrained fixed parameters, we use the normal distribution, $\boldsymbol{\beta}^U \sim N(\mathbf{b}_0, \boldsymbol{B}_0)$, where $\mathbf{b}_0 = \mathbf{0}_{TH-9}$ and $\boldsymbol{B}_0 = 10\boldsymbol{I}_{TH-9}$. For the error precision, we use the gamma distribution, $\sigma_e^{-2} \sim \mathcal{G}(g_1, g_2)$, where $g_1 = 1$ and $g_2 = 1$. Following Chen and Dunson (2003), we let

$$p(\boldsymbol{\lambda}^k, \boldsymbol{\gamma}^k) = p(\boldsymbol{\gamma}^k|\boldsymbol{\lambda}^k)p(\boldsymbol{\lambda}^k), \quad k \in \{\mathcal{H}, \mathcal{T}\},$$

and further assume that

$$p(\boldsymbol{\gamma}^k \mid \boldsymbol{\lambda}^k) \propto N\left(\boldsymbol{\gamma}^k; \boldsymbol{\gamma}_0^k, \boldsymbol{R}_0^k\right) 1\left(\boldsymbol{\gamma}^k \in \mathcal{R}_{\boldsymbol{\lambda}^k}\right),$$

and

$$p(\boldsymbol{\lambda}^k) = \prod_l p\left(\lambda_l^k\right),$$

where $1(\cdot)$ is an indicator function and $\mathscr{R}_{\boldsymbol{\gamma}^k}$ specifies the constraint that whenever an element of $\boldsymbol{\gamma}^k$ is zero, the elements in $\boldsymbol{\gamma}$ that are in the same row number are also zero. Furthermore, we let $p(\lambda_l^k) \overset{d}{=} \text{ZIN}^+(p_{l0}^k, m_{l0}^k, (s_{l0}^k)^2)$, $k \in \{H, T\}$, where $\text{ZIN}^+(\pi, \mu, \sigma^2)$ denotes the density of a zero inflated positive normal distribution with a point mass π at zero and a $N(\mu, \sigma^2)$ density truncated to be in the positive real line. For our analysis, the prior parameters corresponding to the half normal distribution are $p_0 = 0.1$, $m_0 = 0$, and $s_0 = 1$ for each of the four hormone-specific subject random effect variances ($\lambda_E^{\mathscr{H}}, \lambda_L^{\mathscr{H}}, \lambda_F^{\mathscr{H}}, \lambda_P^{\mathscr{H}}$) and the eight visit-specific subject random effect variances $\lambda_1^{\mathscr{T}}, \ldots, \lambda_8^{\mathscr{T}}$. The correlation parameters γ_{lm} are chosen a priori to be $N(0, 1)$ with the added restriction that if the corresponding variance parameter is selected to be zero, then all correlation parameters involving that particular random effect are also set to zero.

The prior distribution for the constrained parameters, $\boldsymbol{\beta}^C$, is very similar to our illustrative example. Each partition of the constrained parameter can be represented as $\boldsymbol{\beta}^{C_h} = \boldsymbol{V}^h\boldsymbol{\alpha}^h + \boldsymbol{U}^h\boldsymbol{\gamma}^h$ for $h \in \{E, L, P\}$. Note that $\boldsymbol{V}^h$ is zero for $h \in \{E, P\}$, therefore we do not need prior distributions for $\boldsymbol{\alpha}^E$ and $\boldsymbol{\alpha}^P$ because they will not affect the distribution of $\boldsymbol{\beta}^C$. Additionally, suppose we use a prior density $\pi_{\boldsymbol{\alpha}^L}$ on $\boldsymbol{\alpha}^L$ and $\pi_{\boldsymbol{\gamma}^h}$ on $\boldsymbol{\gamma}^h$ for $h \in \{E, L, P\}$. We denote the MW prior induced on $\boldsymbol{\beta}^C$ as $p(\boldsymbol{\beta}^C) = p^{\text{MW}}(\pi_{\boldsymbol{\gamma}^E}, \pi_{\boldsymbol{\alpha}^L}, \pi_{\boldsymbol{\gamma}^L}, \pi_{\boldsymbol{\gamma}^P})$. We assume the Dirichlet distribution for $\boldsymbol{\alpha}^L$ and independent gamma distributions for each element in $\boldsymbol{\gamma}^h$, for $h \in \{E, L, P\}$. Since our prior knowledge is limited to knowledge about the constraint space, we choose the parameters of the Dirichlet and gamma distributions to be weakly informative over the constraint space. We explore choices for the Dirichlet parameter by plotting histograms on each dimension of $\boldsymbol{V}^L\boldsymbol{\alpha}^L$ to find a nearly uniform distribution of starting locations on the simplex. We choose Dir(2, 0.4975, 0.005, 0.4975) for the prior of $\boldsymbol{\alpha}^L$, and $\mathscr{G}(1, 2)$ with mean 2 and variance 4 for the prior of each element in $\boldsymbol{\gamma}^h$, for $h \in \{E, L, P\}$. To obtain posterior estimates, we generate 3 independent Gibbs chains with random starting locations of size 30,000 following a burn-in of 20,000.

Figure 7 shows the population mean hormone level estimates that comply with the constraints and have profiles that we would expect from Figure 1. Additionally, a posterior summary of the covariate effects are given in Table 2. We report the mean of the posterior sample and the 90% credible intervals. The last column of Table 2 displays an asterisk for credible intervals that do not include zero. The oxidative stress marker, F_2-isoprostanes, was observed to be significantly positively associated with progesterone, and estrogen levels. These results corroborate findings in Schisterman et al. (2010) and Roy et al. (2011), where estrogen and progesterone levels were significantly positively associated with F_2-isoprostanes. Additionally, these results are supported by Figure 6. Age was also positively associated with FSH levels, a result consistent with previous research showing FSH to be a marker of ovarian reserve, with increased levels of FSH associated with older age (Broekmans et al. 2006). Estrogen and progesterone levels were negatively associated with age of first menstrual period. Race is coded as a zero to indicate Caucasian and one to indicate other racial groups. Therefore, the results indicate that other racial groups have significantly higher levels of estrogen and lower levels of LH than Caucasians. There is a growing body of literature showing that hormone levels differ by race (Goldin et al. 1986; Ursin et al. 2001; Randolph et al. 2003; Santoro et al. 2004).

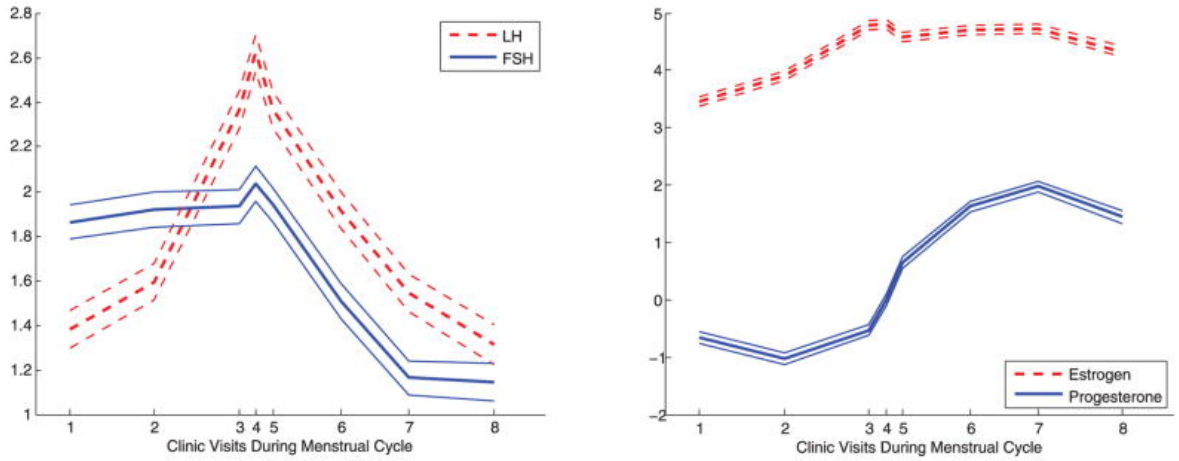


Figure 7

Posterior mean estimates and 90% credible intervals (gonadotropic hormones left panel: FSH and LH; ovarian hormones right panel: estrogen and progesterone). The shape of each curve conforms to the known forms of the hormonal levels. The online version of this figure is in color.

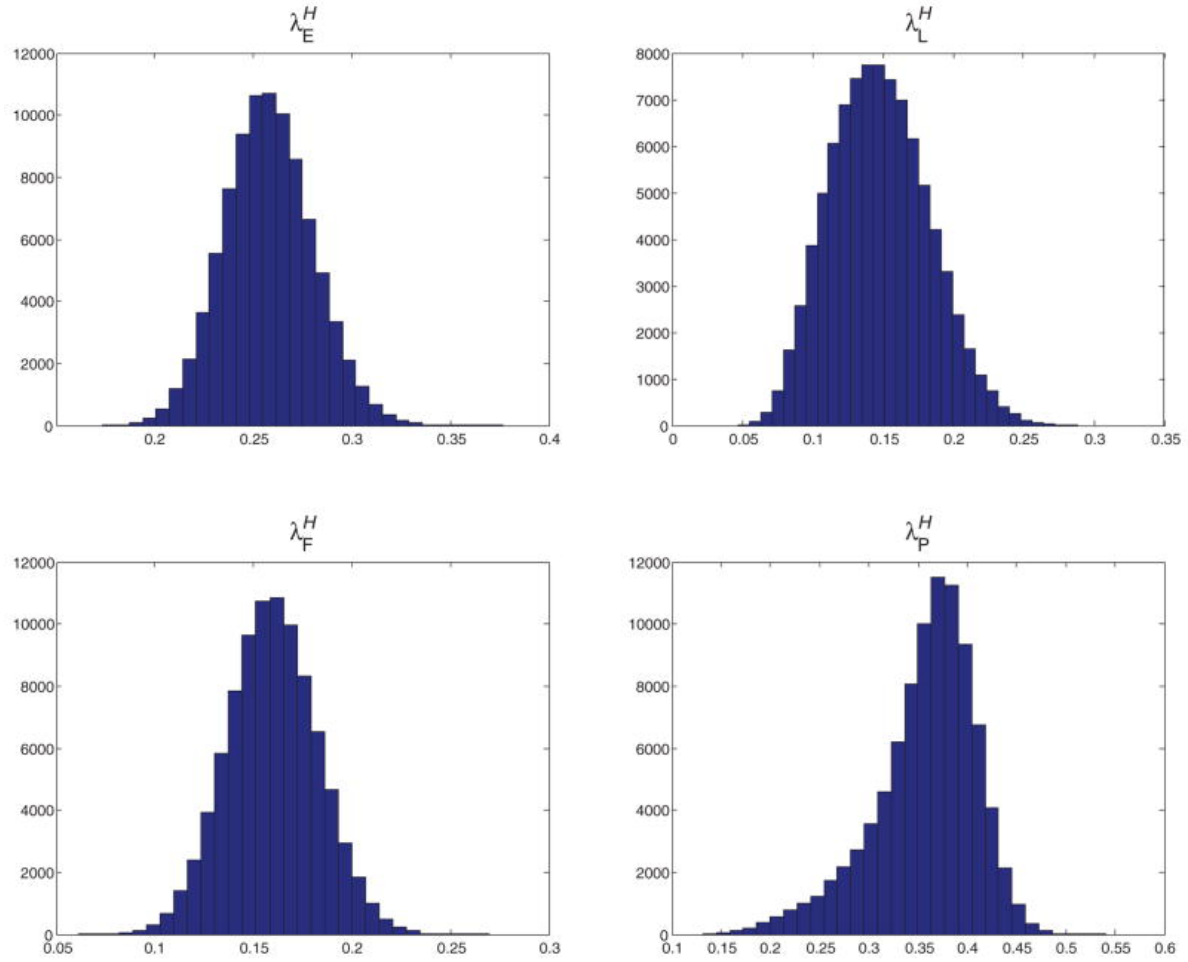
Table 2

Posterior summary of fixed effect estimates using data from the BioCycle Study

	Progesterone	0.002	−0.058, 0.062	
Race	Estrogen	0.083	0.043, 0.124	*
	FSH	−0.020	−0.054, 0.014	
	LH	−0.075	−0.116, −0.034	*
	Progesterone	0.023	−0.036, 0.081	
Age at menarche	Estrogen	−0.043	−0.082, −0.004	*
	FSH	0.013	−0.019, 0.046	
	LH	−0.034	−0.074, 0.006	
	Progesterone	−0.076	−0.135, −0.018	*
Time varying covariate effects on hormones				
log(beta-carotene)	Estrogen	0.000	−0.037, 0.036	
	FSH	0.000	−0.032, 0.032	
	LH	−0.015	−0.052, 0.023	
	Progesterone	0.048	0.001, 0.095	*
log(γ-vitamin)	Estrogen	0.021	−0.014, 0.056	
	FSH	0.012	−0.018, 0.043	
	LH	−0.014	−0.050, 0.021	
	Progesterone	−0.011	−0.053, 0.032	
Cholesterol	Estrogen	0.027	−0.013, 0.066	
	FSH	−0.033	−0.068, 0.001	
	LH	0.003	−0.038, 0.043	
	Progesterone	0.042	−0.010, 0.094	
Homocysteine	Estrogen	−0.049	−0.088, −0.010	*
	FSH	0.038	0.006, 0.070	*
	LH	0.015	−0.025, 0.055	
	Progesterone	−0.002	−0.060, 0.056	
log(F_2-isoprostane)	Estrogen	0.038	0.002, 0.074	*
	FSH	−0.032	−0.064, 0.000	
	LH	−0.013	−0.049, 0.024	
	Progesterone	0.072	0.027, 0.116	*

Open in a separate window

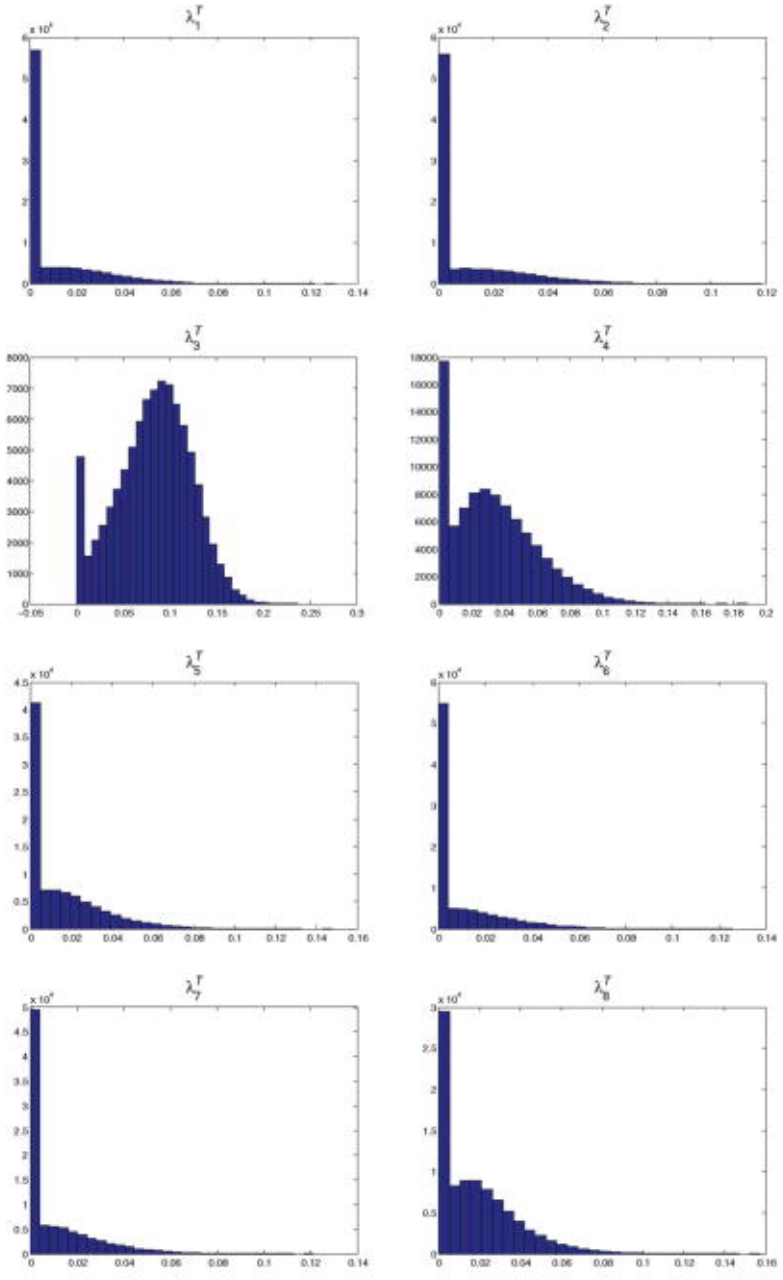
Figure 8 shows the histograms of posterior samples of hormone-specific subject random effects components, ($\lambda_E^{\mathcal{H}}, \lambda_L^{\mathcal{H}}, \lambda_F^{\mathcal{H}}, \lambda_P^{\mathcal{H}}$). The histograms suggest significant between-subject variability across all four hormones. Figure 9 gives the histograms of the visit-specific subject random effect components, ($\lambda_1^{\mathcal{T}}, \ldots, \lambda_8^{\mathcal{T}}$). The subject effect components for all visits except visit 3, and to some extent visits 4 and 8, have a large mass at zero. These subject effects indicate a large variation in the pattern of the LH surge and progesterone rise across women, which was also seen in Roy et al. (2011). Figure 10 shows the fitted versus observed hormone observations for each of the eight visits. These plots indicate that our model is capturing most of the random variability in the hormone measurements. In the beginning of the menstrual cycle, the observed progesterone levels are very low, and the precision for low levels is not as good as it is for higher levels. This is why the observed progesterone levels appear to not be continuous for the first three visits. We observe a similar result for the estrogen levels of the first visit.



Open in a separate window

Figure 8

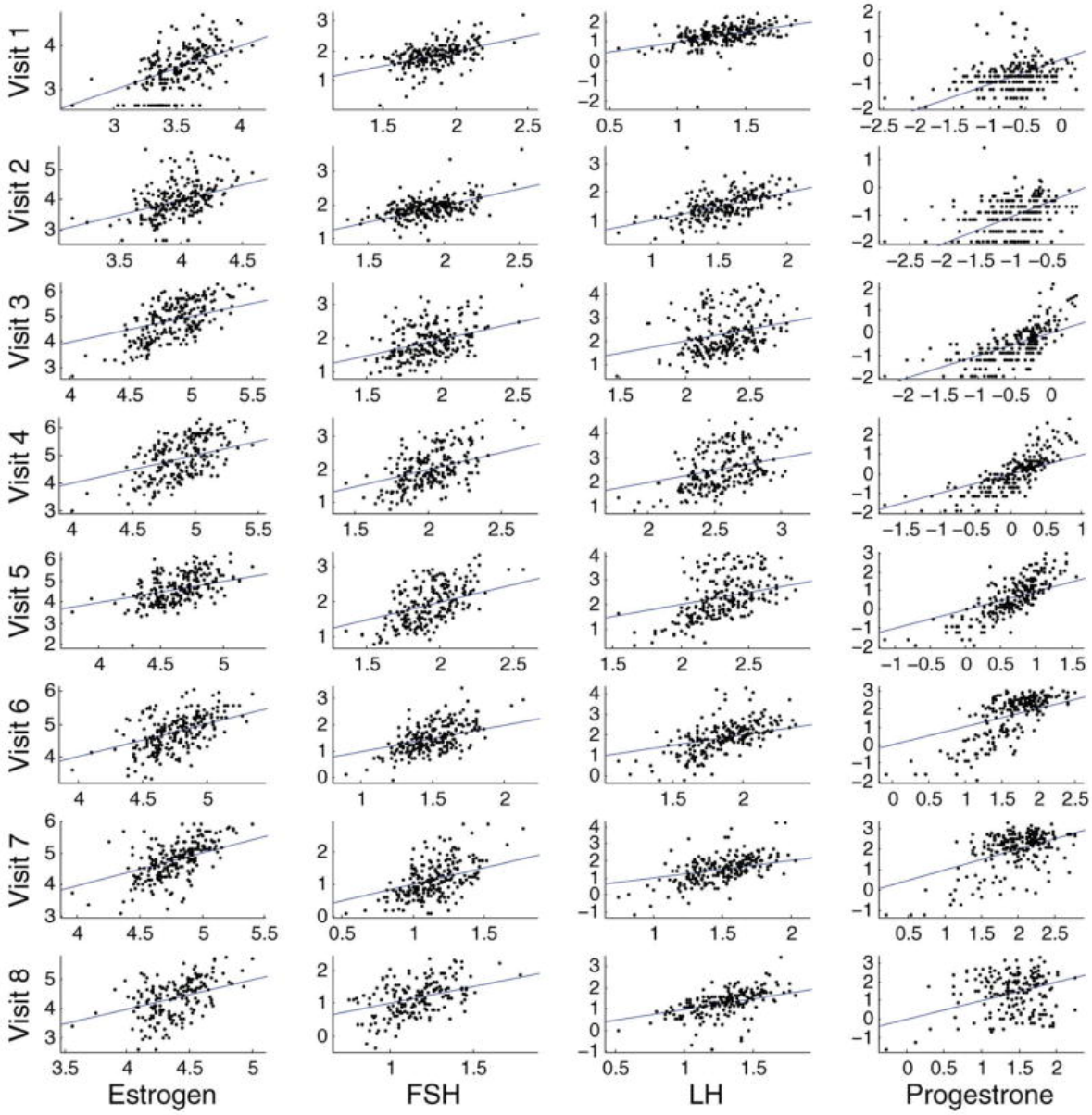
Posterior densities of the hormone-specific subject random effect variances. The online version of this figure is in color.



Open in a separate window

Figure 9

Posterior densities of the visit-specific subject random effect variances. The online version of this figure is in color.



Open in a separate window

Figure 10

Observed versus fitted hormone levels for each visit. Observed hormone levels for each individual are shown on the *x*-axis and the fitted hormone levels shown on the *y*-axis. The plot at the top left corner is the estrogen levels for the first visit, and the plot at the bottom right corner is the progesterone levels for the last visit. The line has slope 1 and intercept 0 and represents a perfect relationship between the observed and fitted hormone levels. The online version of this figure is in color.

4.3 Results and Discussions

We have defined a new class of priors that are flexible and permit full Bayesian analysis for models with linear inequality constraints. The MW priors have been used to analyze the data from the BioCycle Study. Our findings from the analysis of the BioCycle data using the MW prior provide new insight regarding the association between reproductive hormone levels and oxidative stress. The model that we used has the flexibility to answer complicated questions about the variability of hormone levels during the menstrual cycle. The unique sampling design of the BioCycle Study allowed us to treat the visit time as a categorical variable, allowing us to impose constraints among the levels of the hormones at different visits. For continuous time measurement, the MW prior may not be applicable, and a nonparametric Bayesian approach may be more suitable. While we have ignored data from the first cycle due to increased variability, terms related to cycles can be easily incorporated in the model to answer questions about between cycle variability. Understanding between cycle variability is also important to reproductive epidemiologists.

Acknowledgments

This research was partially supported by the Long-Range Research Initiative of the American Chemistry Council and the Intramural Research Program of the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health.

Contributor Information

Michelle R. Danaher, Predoctoral Fellow, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, Rockville, MD 20852.

Anindya Roy, Professor, Department of Mathematics and Statistics, University of Maryland Baltimore County, Baltimore, MD 21250.

Zhen Chen, Investigator, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, Rockville, MD 20852.

Sunni L. Mumford, Research Fellow, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, Rockville, MD 20852.

Enrique F. Schisterman, Chief and Senior Investigator, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, Rockville, MD 20852.

References

1. Broekmans F, Kwee J, Hendriks D, Mol B, Lambalk C. A Systematic Review of Tests Predicting Ovarian Reserve and IVF Outcome. *Human Reproduction.* 2006;12:685–718. [PubMed] [Google Scholar]
2. Chen M, Deely J. Bayesian Analysis for a Constrainted Linear Multiple Regression Problem for Predicting the New Crop of Apples. *Journal of Agricultural, Biological, and Environmental Statistics.* 1996;1:467–489. [Google Scholar]
3. Chen M, Shao Q. Monte Carlo Methods for Bayesian Analysis of Constrained Parameter Problems. *Biometrika.* 1998;85:73–87. [Google Scholar]

4. Chen Z, Dunson D. Random Effects Selection in Linear Mixed Models. *Biometrics.* 2003;59:762–769. [PubMed] [Google Scholar]
5. Chung Y, Rabe-Hesketh S, Gelman A, Lui J, Dorie V. *Avoiding Boundary Estimates in Linear Mixed Models Through Weakly Informative Priors.* 2011 Available at http://www.stat.columbia.edu/gelman/research/unpublished/avoid.pdf.
6. Davis W. Bayesian Analysis of the Linear Model Subject to Linear Inequality Constraints. *Journal of the American Statistical Association.* 1978;73:573–579. [Google Scholar]
7. Fritz M, Speroff L. The Endocrinology of the Menstrual Cycle: The Interaction of Folliculogenesis and Neuroendocrine Mechanisms. *Fertility and Sterility.* 1982;38:509–529. [PubMed] [Google Scholar]
8. Galindo-Garre F, Vermunt J. Avoiding Boundary Estimates in Latent Class Analysis by Bayesian Posterior Mode Estimation. *Behaviormetrika.* 2006;33:43–59. [Google Scholar]
9. Gaskins AJ, Mumford SL, Zhang C, Wactawski-Wende J, Hovey KM, Whitcomb BW, Howards PP, Perkins N, Yeung E, Schisterman EF. Effect of Daily Fiber Intake on Reproductive Function: The Bio-Cycle Study. *The American Journal of Clinical Nutrition.* 2009;90:1061–1069. [PMC free article] [PubMed] [Google Scholar]
10. Gelfand A, Smith A, Lee T. Bayesian Analysis of Constrained Parameters and Truncated Data Problems. *Journal of the American Statistical Association.* 1992;87:523–532. [Google Scholar]
11. Gelman A, Meng X-L, Stern H. Posterior Predictive Assessment of Model Fitness via Realized Discrepancies. *Statistica Sinica.* 1996;6:733–807. [Google Scholar]
12. Geweke J. Exact Inference in the Inequality Constrained Normal Linear Regession Model. *Journal of Applied Econometrics.* 1986;1:127–141. [Google Scholar]
13. Geweke J. Bayesian Inference for Linear Models Subject to Linear Inequality Constraints. In: Lee JC, Johnson WO, Zellner A, editors. *Modeling and Prediction: Honouring Seymour Geisser.* New York: Springer; 1996. pp. 248–263. [Google Scholar]
14. Gilbert S. *Developmental Biology.* 8. Sunderland, MA: Sinauer Associates Inc; 2006. [Google Scholar]
15. Goldin B, Adlercreutz H, Gorbach S, Woods M, Dwyer J, Conlon T, Bohn E, Gershoff S. The Relationship Between Estrogen Levels and Diets of Caucasian American and Oriental Immigrant Women. *The American Journal of Clinical Nutrition.* 1986;44:945–953. [PubMed] [Google Scholar]
16. Gunn L, Dunson D. A Transformation Approach for Incorporating Monotone or Unimodal Constraints. *Biostatistics.* 2005;6:434–449. [PubMed] [Google Scholar]
17. Hall J. Neuroendocrine Control of the Menstrual Cycle. In: Strauss JF III, Barbieri RL, editors. *Yen and Jaffe's Reproductive Endocrinology.* 6. Philadelphia, PA: Elsevier; 2009. pp. 139–154. [Google Scholar]
18. Hoffman K. Generalized Bayes Stein-Type Estimators for Regression Parameters Under Linear Constraints. *Journal of Multivariate Analysis.* 1993;46:120–130. [Google Scholar]
19. Howards P, Schisterman E, Wactawski-Wende J, Reschke J, Frazer A, Hovey K. Timing Clinic Visits to Phases of the Menstrual Cycle by Using a Fertility Monitor: The BioCycle Study. *American Journal of Epidemiology.* 2009;169:105–112. [PMC free article] [PubMed] [Google Scholar]

20. Kato B, Hoijtink H. A Bayesian Approach to Inequality Constrained Linear Mixed Models: Estimation and Model Selection. *Statistical Modelling.* 2006;6:231–249. [Google Scholar]
21. Klugkist I, Laudy O, Hoijtink H. Bayesian Eggs and Bayesian Omelettes: Reply to Stern. *Psychological Methods.* 2005a;10:500–503. [PubMed] [Google Scholar]
22. Klugkist I, Laudy O, Hoijtink H. Inequality Constrained Analysis of Variance: A Bayesian Approach. *Psychological Methods.* 2005b;10:477–493. [PubMed] [Google Scholar]
23. Leonard T, Hsu J. Bayesian Inference for a Covariance Matrix. *The Annals of Statistics.* 1992;20:1669–1696. [Google Scholar]
24. Mandelkern M. Setting Confidence Intervals for Bounded Parameters. *Statistical Science.* 2002;17:149–172. [Google Scholar]
25. Marchand E, MacGibbon B. Minimax Estimation of a Constrained Binomial Proportion. *Statistical Decisions.* 2000;18:129–167. [Google Scholar]
26. Marchand E, Perror F. Improving on the MLE of a Bounded Normal Mean. *The Annals of Statistics.* 2001;29:1078–1093. [Google Scholar]
27. Mulder J, Hoijtink H, Klugkist I. Equality and Inequality Constrained Multivariate Linear Models: Objective Model Selection Using Constrained Posterior Priors. *Journal of Statistical Planning and Inference.* 2010;140:887–906. [Google Scholar]
28. Mulder J, Klugkist I, van de Schoot R, Meeusb W, Selfhout M, Hoijtink H. Bayesian Model Selection of Informative Hypotheses for Repeated Measurements. *Journal of Mathematical Psychology.* 2009;53:530–546. [Google Scholar]
29. Neuman E. *Tutorial 6: Linear Programming With Matlab.* n.d [online], Available at http://www.math.siu.edu/matlab/tutorial6.pdf.
30. Pindyck R, Rubinfeld D. *Econometric Models and Economic Forecasts.* New York: McGraw-Hill; 1981. [Google Scholar]
31. Randolph J, Sowers M, Gold E, Mohr B, Luborsky J, Santoro N, Mc-Connell D, Finkelstein J, Korenman S, Matthews K, Sternfeld B, Lasley B. Reproductive Hormones in the Early Menopausal Transition: Relationship to Ethnicity, Body Size, and Menopausal Status. *The Journal of Clinical Endocrinology and Metabolism.* 2003;88:1516–1522. [PubMed] [Google Scholar]
32. Robertson T, Wright F, Dykstra R. *Order Restricted Statistical Inference.* New York: Wiley; 1988. [Google Scholar]
33. Roy A, Danaher M, Mumford S, Chen Z. A Bayesian Order Restricted Model for Hormonal Dynamics During Menstrual Cycles of Healthy Women. *Statistics in Medicine.* 2011 doi: 10.1002/sim.4419. [PubMed] [CrossRef] [Google Scholar]
34. Santoro N, Lasley B, McConnell D, Allsworth J, Crawford S, Gold E, Finkelstein J, Greendale G, Kelsey J, Korenman S, Luborsky J, Matthews K, Midgley R, Powell L, Sabatine J, Schocken M, Sowers M, Weiss G. Body Size and Ethnicity are Associated With Menstrual Cycle Alterations in Women in the Early Menopausal Transition: The Study of Women's Health Across the Nation (SWAN) Daily Hormone Study. *The Journal of Clinical Endocrinology and Metabolism.* 2004;89:2622–2631. [PubMed] [Google Scholar]
35. Schisterman E, Gaskins A, Mumford S, Browne R, Yeung E, Trevisan M, Hediger M, Zhang C, Perkins N, Hovey K, Wactawski-Wende J. Influence of Endogenous Reproductive Hormones on f_2-isoprostane Levels in Premenopausal Women. *American Journal of Epidemiology.*

2010;173(4):430–439. [PMC free article] [PubMed] [Google Scholar]
36. Scolaro KL, Lloyd KB, Helms KL. Devices for Home Evaluation of Women's Health Concerns. *American Journal of Health-System Pharmacy.* 2008;65:299–314. [PubMed] [Google Scholar]
37. Silvapulle M, Sen P. *Constrained Statistical Inference: Inequality, Order, and Shape Restrictions.* Hoboken, NJ: Wiley; 2004. [Google Scholar]
38. Stern H. Model Inference or Model Selection: Discussion of Klugkist, Laudy and Hoijtink. *Psychological Methods.* 2005;10:494–499. [PubMed] [Google Scholar]
39. Ursin G, Wilson M, Henderson B, Kolonel L, Monroe K, Lee H, Seow A, Yu M, Stanczyk F, Gentzschein E. Do Urinary Estrogen Metabolites Reflect the Differences in Breast Cancer Risk Between Singapore Chinese and United States African-American and White Women? *Cancer Research.* 2001;61:3326–3329. [PubMed] [Google Scholar]
40. van Dyk D. Comment on 'Setting Confidence Intervals for Bounded Parameters' by Mark Mandelkern. *Statistical Science.* 2002;17:164–168. [Google Scholar]
41. Wactawski-Wende J, Schisterman E, Hovey K, Howards P, Browne R, Hediger M, Liu A, Trevisan M. BioCycle Study: Design of the Longitudinal Study of the Oxidative Stress and Hormone Variation During the Menstrual Cycle. *Paediatric and Perinatal Epidemiology.* 2009;23(2):171–184. [PMC free article] [PubMed] [Google Scholar]