A Simulation Technique for GMRFs

Jason K. Johnson

April 17, 2001

Abstract. *This brief considers a simple technique for the efficient Monte-Carlo generation of sample paths of a Gauss-Markov random field. The approach exploits the Markov structure of the field to instead simulate the conjugate process in a straight-forward manner (at least when the inverse covariance is diagonally dominant, but extensions of this method might apply otherwise). Sample paths of the original field may then be obtained by executing an existing optimal estimation algorithm or by exploiting existing iterative methods for solving linear systems which are both sparse and symmetric.*

Gauss-Markov Random Fields (GMRFs) are jointly Gaussian processes which admit compact description as graphical models. Consider such a process

$$\mathrm{x} \sim N(\bar{x}, P) \tag{1}$$

having Markov structure characterized by the graph $\mathcal{G} = (\mathcal{S}, \mathcal{E})$ where $\mathcal{S}$ is the set of vertices of the graph (identified with the set of sites of the field indexing the elements of x) and $\mathcal{E}$ is the set of all (undirected) edges of the graph (representing interactions between sites within the field). An edge between sites s and t is denoted $\langle st \rangle$ (which is not distinguished from $\langle ts \rangle$). Let J denote the information matrix P^{-1}.

An important result from the theory of GMRFs (see [3] and [4]) is that the interaction between sites s and t as measured by the *partial correlation coefficient* between states x_s and x_t given the state of the rest of the field x^c_{st} is directly related to the corresponding element of the information matrix $J_{s,t}$ by

$$\rho(\mathrm{x}_s, \mathrm{x}_t | \mathrm{x}^c_{st}) = -\frac{J_{s,t}}{\sqrt{J_{s,s} J_{t,t}}}. \tag{2}$$

This gives the well-known result ([1], [3] and [4]) that states x_s and x_t are conditionally independent given $x^c_{s,t}$ if and only if $J_{s,t} = 0$.

$$(\mathrm{x}_s \perp \mathrm{x}_t \mid \mathrm{x}^c_{st}) \iff J_{s,t} = 0 \tag{3}$$

Assuming $\mathcal{G}$ is minimal in the sense that it has as few edges as possible so as to capture all such independence relations (this assumption is not problematic for Gaussian processes) then the sparsity of the information matrix agrees with the structure of $\mathcal{G}$.

$$J_{s,t} = 0 \iff \langle st \rangle \in \mathcal{E} \tag{4}$$

We may exploit the sparsity of the information matrix to simulate GMRFs as follows. Consider the Gaussian random vector

$$\mathrm{h} = P^{-1}\mathrm{x} \sim N(\bar{h}, J) \tag{5}$$

where $\bar{h} = P^{-1}\bar{x}$. The processes x and h are referred to as *conjugate processes* since the covariances of the two processes are inverses of one another. Consider the relation of the conjugate process h to the graph $\mathcal{G}$. Since $\mathrm{cov}(\mathrm{h}) = J$ then by (4) we see that h_s and h_t are dependent ($\mathrm{cov}(\mathrm{h}_s, \mathrm{h}_t) \neq 0$) if and only if s and t are adjacent in $\mathcal{G}$. Or, equivalently, h_s and h_t are independent if and only if they are not adjacent in $\mathcal{G}$.

$$\mathrm{h}_s \perp \mathrm{h}_t \iff \langle st \rangle \notin \mathcal{E} \tag{6}$$

This suggests the following realization of process h. For each edge $\langle st \rangle \in \mathcal{E}$ simulate an independent sample of the standard normal distribution $N(0, 1)$ and denote this deviate $w_{\langle st \rangle}$. Likewise, for each site $s \in \mathcal{S}$ simulate a standard normal deviate v_s. Given these deviates, form the sample path h' by

$$h'_s = \bar{h}_s + \alpha_s v_s + \sum_{\langle st \rangle} \beta_{s,t} w_{\langle st \rangle} \tag{7}$$

where the sum is taken with s fixed and t varying over all sites adjacent to s (such that $\langle st \rangle \in \mathcal{E}$). Note that there are two parameters $(\beta_{s,t}, \beta_{t,s})$ for each edge $\langle st \rangle \in \mathcal{E}$. The complexity of simulating one such sample-path is $|\mathcal{S}| + |\mathcal{E}|$ both in terms of the number of required Gaussian deviates and the number floating-point multiply-and-add operations required to form the sample path.

Clearly, h' is a Gaussian process with mean $\bar{h}$. Let us compute the covariance of h' to see if we can choose the realization parameters $(\{\alpha_s\}, \{\beta_{s,t}\})$ so as to match the covariance of h. Since the v and

w deviates are mutually independent and have unit-variance, the variance at site s is

$$\mathrm{var}(\mathrm{h}'_s) = \alpha_s^2 + \sum_{\langle st \rangle} \beta_{s,t}^2. \tag{8}$$

Since each element of h' only depends on those edge-deviates defined on adjacent edges we have (for $s \neq t$) that

$$\mathrm{cov}(\mathrm{h}'_s, \mathrm{h}'_t) = \begin{cases} \beta_{s,t}\beta_{t,s}, & \langle st \rangle \in \mathcal{E} \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

which has precisely the same sparsity structure as J. If we choose the realization parameters to satisfy

$$\beta_{s,t}\beta_{t,s} = J_{s,t} \ \ \forall \langle st \rangle \in \mathcal{E} \tag{10}$$

and

$$\alpha_s^2 + \sum_{\langle st \rangle} \beta_{s,t}^2 = J_{s,s} \ \ \forall s \in \mathcal{S} \tag{11}$$

then we will have constructed a realization of h. Of coarse, it is necessary (sufficient?) that J be non-negative for such a realization to exist.

In the case that J is (marginally) diagonally-dominant such that

$$J_{s,s} \geq \sum_{t \neq s} |J_{s,t}| \ \ \forall s \in \mathcal{S} \tag{12}$$

then we may satisfy these constraints as follows. For all ordered pairs (s,t) such that $\langle st \rangle \in \mathcal{E}$ let

$$\beta_{s,t} = \begin{cases} \sqrt{J_{s,t}}, & s < t \\ J_{s,t}/\sqrt{J_{s,t}}, & s > t. \end{cases} \tag{13}$$

For all sites s let

$$\alpha_s = \left(J_{s,s} - \sum_{\langle st \rangle} |J_{s,t}| \right)^{1/2}. \tag{14}$$

This construction works because the argument of the square root in (14) is non-negative by (12). These realization parameters then satisfy the constraints (10) and (11) so as to realize h. The complexity of constructing this realization is dominated by the $|\mathcal{S}|+|\mathcal{E}|$ required square-root computations.

Note. If J is not diagonally dominant, then we might instead consider some sort of iterative constraint propagation algorithm to identify a realization of h. For instance, this might be posed as the minimization problem $\min_{\alpha,\beta} \|J - J'\|_F^2$ where J' is the covariance of h$'$ and $\|\cdot\|_F$ denotes the Frobenius matrix norm. Apparently this minimization may be implemented by a local fixed-point iteration provided we enforce the inequality constraints

$$\sum_{\langle st \rangle} \beta_{s,t}^2 \leq J_{s,s} \ \ \forall s \in \mathcal{S} \tag{15}$$

which does not result in any loss of generality since by (8) any realization of h must satisfy (15). However, I have not implemented such an algorithm and will not discuss it further here.

In any case, the identification of a realization of h need only be performed once per model rather then once per sample-path. This observation may be useful when performing Monte-Carlo performance estimation with respect to a fixed model.

Once we have a sample path h' we may then construct a sample path x' of the original random field x by solving the sparse system of linear equations

$$Jx' = h'. \tag{16}$$

Such symmetric systems may be efficiently solved using standard iterative methods (see, for instance [2]) such as preconditioned conjugate gradients perhaps employing an incomplete Cholesky factorization of J to implement the preconditioner.

Note also, that the solution of the system (16) also arises when the optimal (MMSE) estimate of x is required given the graphical model $(\bar{h}, J)$. This is equivalent to solving $J\bar{x} = \bar{h}$. Hence, if we have a specialized inference/estimation algorithm for solving this latter problem (either exactly or approximately) then we might exploit this algorithm to also perform (exact or approximate) simulation simply by replacing $\bar{h}$ by h' in our graphical model and identifying the output "estimate" as a sample path of x. For instance, loopy belief propagation might be used to perform simulation of GMRFs in this manner.

References

[1] Frakt, A.B. (1999). *Internal Multiscale Autoregressive Processes, Stochastic Realization, and Covariance Extension.* PhD thesis. Laboratory for Information and Decision Systems, MIT.

[2] Golub, G.H. and Van Loan, C.F. (1996). *Matrix Computations, 3rd ed.* The Johns Hopkins University Press.

[3] Lauritzen, S.L. (1996). *Graphical Models.* Oxford Science Publications.

[4] Guyon, X. (1995). *Random Fields on a Network: modeling, statistics, and applications.* Springer-Verlag.