Assignment 2, due February 26

Corrections: (none yet).

1. (*"Gaussian" in probability plays the role of "linear" in other parts of mathematics. You try to understand the Gaussian case of anything you do. Gaussians are important in MCMC for several reasons. We sometimes use MCMC even for Gaussians if the dimension is so high that the direct Choleski method is impractical. There are many MCMC methods whose behavior can be understood in the Gaussian case.*) A linear Gaussian recurrence relation takes the form

$$X_{n+1} = AX_n + BZ_n \; , \tag{1}$$

where $X_n \in \mathbb{R}^d$, and A is a $d \times d$ matrix, and $Z_n \sim \mathcal{N}(0, \mathbb{I}_{d\times d})$, and B is another $d \times d$ matrix. The Z_n are i.i.d. Let $\mu_1, \ldots, \mu_d$ be the eigenvalues of A. The iteration (1) is *stable* if $|\mu_j| < 1$ for $j = 1, \ldots, d$. The iteration is *non-degenerate* if $\mu_j \neq 0$ for $|\mu_j| < 1$ for $j = 1, \ldots, d$. This exercise shows that for any initial distribution, the distribution of X_n converges to the unique invariant distribution, which is Gaussian.

If $X_0 = 0$, then X_n is a linear combination of mean zero Gaussians, so it is a mean zero Gaussian. Let $C_n = E[X_n X_n^t]$ be the covariance matrix of X_n.

(a) Show that
$$C_{n+1} = AC_nA^t + BB^t \; . \tag{2}$$

(b) Let $\mathcal{L}$ be the vector space of real $d \times d$ symmetric matrices. This has dimension $d(d+1)/2$. Show that the map (2) may be written $C_{n+1} = L_A C_n + f_B$, where L_A is a linear transformation on $\mathcal{L}$, and $f_B \in \mathcal{L}$, and $C_n \in \mathcal{L}$. You can think of L_A as a $d(d-1)/2 \times d(d-1)/2$ matrix built from A, and C_n as a column vector with the $d(d-1)/2$ entries of C_n arranged in some order. Of course, f_B is a column vector whose entries entries of the matrix BB^t.

(c) Let M be any square matrix. The *spectral radius* of M is $\rho(M) = \max |\lambda|$, where λ is an eigenvalue of M. Show that the spectral radius of $\rho(L_A) = \rho(A)^2$. There are several ways to do this. Pick one.

Hint (1): If $Ar_j = \mu_j r_j$ then $r_j r_k^t$ is an "eigenmatrix" of L_A.

Hint (2): What is L_A^n? Equivalently, express C_n in terms of C_0. There is a theorem of Lyapounov that says that if $\rho(A) = r$ and

ϵ is given, then there is a symmetric positive definite matrix W so that if $y = Ax$, then

$$\|y\|_W = \left(y^t W y\right)^{1/2} \leq (r + \epsilon) \|x\|_W \ .$$

(d) Conclude that if A is stable ($\rho(A) < 1$), then $C_n \to C$ as $n \to \infty$, where C is the unique solution of the *Lyapounov* equation $C = ACA^t + BB^t$. Show that the C that satisfies the Lyapounov equation is positive definite if A is stable and non-singular and B is non-singular. (Hint: start with a C_0 that is positive definite in (2)).

(e) let $f_n(x)$ be the probability density of X_n. Let $g(x) = \mathcal{N}(0, C)$ be the multivariate Gaussian density with mean zero and covariance C. Show that this is the unique Gaussian invariant distribution for (1). Show that if $X_0 = 0$, then $f_n \to g$ as $n \to \infty$. This does not yet imply that $f_n \to g$ if f_0 is not Gaussian, or that g is the unique invariant distribution.

(f) Fix the defect of part (e) by looking at iterates X_n for $X_0 \neq 0$. Let $\overline{X}_n$ satisfy (1) with $\overline{X}_0 = 0$ and the same Z_n. Show that $X_n - \overline{X}_n \to 0$ exponentially. Let $\overline{f}_n(x)$ be the density of $\overline{X}_n$. Use this to show that $f_n \to g$ as $n \to \infty$. Hint: (mathematical technicality) if $V(x)$ is a Lipschitz function with compact support, then $E[V(X_n)] - E[V(\overline{X}_n)]$ converges to zero exponentially.

2. (*Many problems in statistical physics involve discretized functions or fields. In such an object, the random object U is a function with values given on a discrete lattice. The phase separation model in the notes is like this. There also are Gaussian fields, which are often called* free fields.) Consider an $n \times n$ lattice with lattice variables $U_{ij} \in \mathbb{R}$ for each (i, j) in the lattice. The energy function is

$$\phi(u) = \frac{1}{2} \sum_{nn} \left(u_{ij} - u_{i'j'}\right)^2 \ . \tag{3}$$

The sum is over nearest neighbor pairs, as described in the Week 3 notes. This exercise discusses MCMC samplers for the free field probability density

$$f(u) = \frac{1}{Z} e^{-\phi(u)/k_B T} \ . \tag{4}$$

Distributions like this are used in Bayesian image reconstruction algorithms and (unfortunately in 4D) to compute the masses of protons and the supposed Higgs boson.

(a) Show that $\phi(u) = \frac{1}{2} u^t M u$, where M is the matrix of the discrete Laplace operator. The boundary conditions for A depend on the boundary conditions for u. For example, if we take $u_{ij} = 0$ on the boundary of the lattice, we get the discrete Laplace operator with Dirichlet boundary conditions. If we omit the terms in (3) with one site on the boundary, we get Neumann boundary conditions.

(b) Consider a mapping $GS_{ij} : u \to \overline{u}$ defined as follows. Hold all of u fixed except for the value at the lattice site (i, j). Minimize the energy over the lattice site variable u_{ij} and use the minimizing value for $\overline{u}_{ij}$. That is, $\overline{u}_{kl} = u_{kl}$ if $(k, l) \neq (i, j)$, and

$$\overline{u}_{ij} = \arg\min_{u_{ij}} \phi(u) \ . \tag{5}$$

Show that this is given by

$$\overline{u}_{ij} = \frac{1}{4} \sum_{nn} u_{i'j'} \ . \tag{6}$$

(This may have to be interpreted properly if (i, j) is on the boundary of the lattice.)

(c) Consider the algorithm (here, `GS(i,j,u)` is the code that implements $GS_{ij}u$.)

```
for ( i = 0; i < n; i++){
   for ( j = 0; j < n; j++) {
      u = GS(i,j,u)
     }
  }
```

This is called the *Gauss Seidel* iteration and is denoted by $u \to Au$. It is an ancient method for solving discrete Laplace equations. Show that the Gauss Seidel iteration $u_{n+1} = A\,u_n$ has $u_n \to 0$ as $n \to \infty$, and that A is a stable matrix. Hint: $\phi(u_{n+1}) < \phi(u_n)$ if $u_n \neq 0$. Warning: do not try to get a quantitative convergence rate. The spectral radius of A is $1 - O(n^{-2})$, so the iterates converge slowly.

(d) Consider the *single site heat bath* MCMC method for sampling (4)

```
for ( i = 0; i < n; i++){
   for ( j = 0; j < n; j++) {
      U = HB(i,j,U)
     }
  }
```

Here `HB(i,j,U)` resamples U_{ij} according to (4) keeping all the other components of U fixed. Show that this has the effect

$$U \ \to \ AU + BZ \ ,$$

where A is the Gauss Seidel matrix and Z is an n^2 component standard normal. Conclude that the single site Gauss Seidel algorithm converges to the correct distribution.

3. (*Although we use detailed balance to derive MCMC samplers, many such samplers do not satisfy detailed balance. Here is an example.*) We want

to sample the discrete three state distribution $P(X = 1) = P(X = 2) = P(X = 3) = \frac{1}{3}$. Consider the MCMC *move* $1 \leftrightarrow 2$ with probability $\frac{1}{2}$. That is, we toss a coin. With probability $\frac{1}{2}$ we do a move and otherwise do nothing. The move we do is $1 \to 2$, $2 \to 1$ and $3 \to 3$. Let R_1 be the 3×3 transition matrix for this move. Similarly, let R_2 be the transition matrix for the $2 \leftrightarrow 3$ move with probability $\frac{1}{2}$. Show that R_1 and R_2 satisfy detailed balance. Show that $R_1 R_2$ preserves the distribution P. You are not allowed to do this by explicit calculation. Show that $R_1 R_2$ does not satisfy detailed balance.