Assignment 4

Problem 1 Rejection sampling

Given only a fair coin (no `random()` function), show how to sample from a weighted coin with 1/3 probability of heads.

Problem 2

Propose a way to sample from a Beta$(1.5, 1.5)$ distribution.

Problem 3

Propose a way to sample from a Naive Bayes model where the features (and parameters) are observed.

Problem 4 Textbook problem 24.2

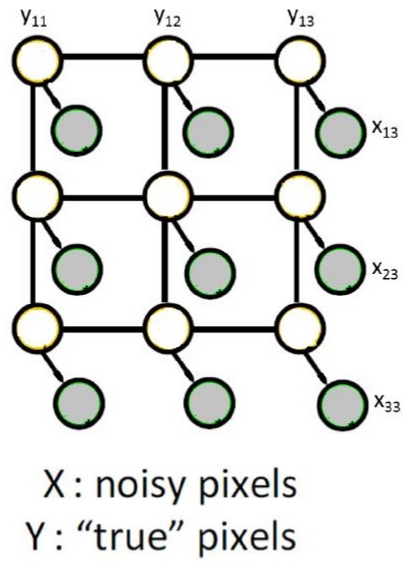
Consider applying Gibbs sampling to a univariate mixture of Gaussians, as in section 24.2.3. Derive the expressions for the full conditionals of π, μ and z. You can use a uniform prior for π and z and a Gaussian prior for μ.

Hint: For π and μ, you might find it hard to derive the normalization constants from scratch. Instead, show that they equal some known distribution (Dirichlet and Gaussian respectively). That is enough to tell you what the normalization constant is – why?

Problem 5 Gibbs sampling

In an image restoration problem, you are given an image corrupted by noise X and you want to recover the original image Y. You can apply Gibbs sampling to a simple Ising Markov random field model to solve this task.

Assignment 4



Graphical model structure, for $N = M = 3$.

Let $x = \{x_{ij}\}$ denote the observed image, with $x_{ij} \in \{-1, +1\}$ representing the pixel at row i and column j. Assume a black-and-white image, with -1 corresponding to white and +1 to black. The image has dimensions $N \times M$, so that $1 \leq i \leq N$ and $1 \leq j \leq M$. Assume a set of (unobserved) variables $\mathbf{y} = \{y_{ij}\}$ representing the true (unknown) image, with $y_{ij} \in \{-1, +1\}$indicating the value of x_{ij} before noise was added. Each (internal) y_{ij} is linked with four immediate neighbors,$y_{i-1,j}$, $y_{i+1,j}$, $y_{i,j-1}$, and $y_{i,j+1}$ which together are denoted $y_{N(i;j)}$. Pixels at the borders of the image (with $i \in \{1, N\}$ or $j \in \{1, M\}$) also have neighbors denoted $y_{N(i;j)}$, but these sets are reduced in the obvious way. We denote E the corresponding set of edges. For example, the pair $((1; 1); (1; 2)) \in E$, but the pair $((1; 1); (2; 2)) \notin E$. The joint probability of y and x can be written (with no prior preference for black or white):

$$p(\mathbf{y}, \mathbf{x}) = \frac{1}{Z}\left\{\prod_{i=1}^{N}\prod_{j=1}^{M} exp^{\eta y_{ij} x_{ij}}\right\} \times \left\{\prod_{((i,j),(i',j'))\in E} exp^{\beta y_{ij} y_{i'j'}}\right\}$$

$$= \frac{1}{Z} exp\left\{\eta \sum_{i=1}^{N}\sum_{j=1}^{M} y_{ij}x_{ij} + \beta \sum_{((i,j),(i',j'))\in E} y_{ij}y_{i'j'}\right\}$$

where

$$Z = \sum_{\mathbf{y},\mathbf{x}} exp\left\{\eta \sum_{i=1}^{N}\sum_{j=1}^{M} y_{ij}x_{ij} + \beta \sum_{((i,j),(i',j'))\in E} y_{ij}y_{i'j'}\right\}$$

(Notice in particular that each pair of neighbors, y_{ij} and $y_{i'j'}$, factors into the formula only once, despite that each variable is a neighbor of the other. Failing to account for this will lead to double counting of β values.) This is equivalent to a Boltzmann (sometimes called Gibbs) distribution with "energy":

$$E(\mathbf{y}, \mathbf{x}) = -\eta \sum_{i=1}^{N} \sum_{j=1}^{M} y_{ij} x_{ij} - \beta \sum_{((i,j),(i',j')) \in E} y_{ij} y_{i'j'}$$

The system will have lower energy, and hence higher probability, in states in which neighboring y_{ij} variables, and neighboring y_{ij} and x_{ij} variables, tend to have the same value (assuming η and β are positive). This captures the fact that each noisy pixel x_{ij} is likely to be similar to the corresponding "true" pixel y_{ij} , and that images tend to be "smooth".

- Derive an expression for the conditional probability that y_{ij} is black given its Markov blanket, i.e. $p(y_{ij} = 1|y_{M(i,j)})$, where $y_{M(i,j)})$ denotes the variables in the Markov blanket of $y_i j$ (but you should be explicit about which variables are included). Your expression should take theform of a logistic function and should depend only on η, β;and $y_{M(i,j)})$.

Problem 6

Show that Gibbs Sampling is a special case of Metropolis-Hastings.