CMPE 548
Monte Carlo methods
Assignment 3

Instructor: A. Taylan Cemgil
TA: Ömer Deniz Akyıldız

Due to April 30, 2014*

1 Change Point Model

Consider the following *coal mining* dataset. The table contains the number of deadly accidents in coal mines occurring between 1851 and 1962 in the UK.

```
[4 5 4 1 0 4 3 4 0 6 3 3 4 0 2 6 3 3 5 4 5 3 1 4 4 1 5 5 3 4 2 5 2 2 3 4 2 1 3 2
 2 1 1 1 1 3 0 0 1 0 1 1 0 0 3 1 0 3 2 2 0 1 1 1 0 1 0 1 0 0 0 2 1 0 0 0 1 1 0 2
 3 3 1 1 2 1 1 1 1 2 4 2 0 0 0 1 4 0 0 0 1 0 0 0 0 0 1 0 0 1 0 1]
```

1. Write down an appropriate model for this problem (consider Poisson intensity change point model covered in the lecture notes.).
2. Using the Gibbs sampler, find the position when the intensity has changed. Can you detect the year when new health and safety regulations are introduced in the late 18th century?

2 Denoising of binary images

2.1 A little bit about inverse problems

This subsection gives introductory knowledge about the image denoising problem for those who did not take 58K. We present a toy example and show how to use the Bayes' theorem for an inverse problem scenario.

Suppose we have a noisy observation $y \in \mathbb{R}$. We denote the ideal, noise-free object with x. A typical inverse problem is that the recovering x from a noisy observation y. The *observation model* can be formulated as,

$$y = x + \eta \tag{1}$$

Assume $\eta \sim \mathcal{N}(\eta; 0, \sigma)$. This model can be equivalently written as,

$$p(y|x) = \mathcal{N}(y; x, \sigma) \tag{2}$$

*Hard-copy submission.

Intuivitely, if you add a certain quantity to a random variable η, the mean increases, if η is Gaussian. This model expresses this fact.

Since we are interested in x, we are interested in something like $p(x|y)$. This is the posterior distribution of x given the data y. This distribution represents the probability distribution of our object of interest. To arrive this quantity, we can use Bayes' theorem,

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)} \tag{3}$$

We can write $p(x, y) = p(y|x)p(x)$. $p(x, y)$ is called joint probability density. In the next problem, our main aim is to draw samples from the posterior distribution by using the Gibbs sampler.

2.2 Image denoising application

Assume we have a noisy (observed) image $Y \in \mathbb{R}^{m \times n}$. So Y is an $m \times n$ matrix. Here Y_{ij} denotes the $Y(i, j)$'th entry of the matrix Y and naturally $Y_{ij} \in \mathbb{R}$. We also suppose the original, noise-free image is a binary image and its pixels are defined as $X_{ij} \in \{-1, 1\}$. We would like to write down a probabilistic model for a class of images and perform inference via the Gibbs sampler. Assume that our observed pixels are corrupted by Gaussian noise. Formally,

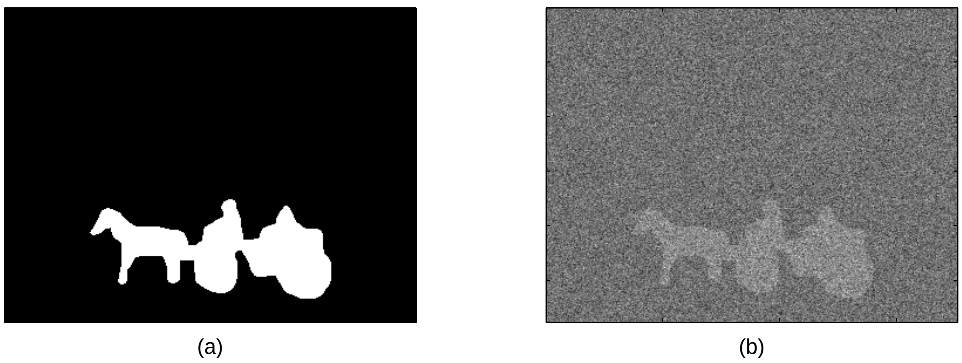

Figure 1: (a) Original image X. (b) Noisy image Y.

$$Y_{ij} = X_{ij} + \eta$$

where $\eta \sim \mathcal{N}(\eta; 0, \sigma)$. As we explained in the previous section, this observation model can equivalently be represented as,

$$p(Y_{ij}|X_{ij}) = \mathcal{N}(Y_{ij}; X_{ij}, \sigma) \tag{4}$$

To recover the image, we define a statistical model of the image. We consider each pixel as a random variable and think of the image as a large graphical model with $m \times n$ nodes. Then, we use a smoothness prior between pixel values. Intuitively, we assume that image is smooth. For a graphical representation of a single pixel and its neighbours see Fig. 2.

We would like to draw samples from the posterior of the each pixel $p(X_{ij}|X_{-ij}, Y_{ij})$[1]. That is, we either draw 1 or -1 using Gibbs sampler since $X_{ij} \in \{-1, 1\}$.

[1] X_{-ij} denotes the all other nodes in the graph (all other pixels).

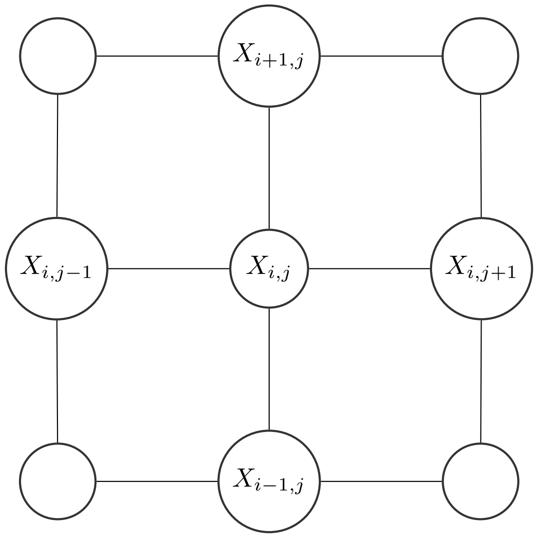

Figure 2: Graphical representation of our model. This model is an instance of a Markov random field.

We define a statistical model (smoothness prior) of pixel relationships as follows.

$$p(X_{ij}, X_{-ij}) = \frac{1}{Z} \exp(J X_{ij} W_{ij}) \tag{5}$$

where

$$\begin{aligned} W_{ij} &= \sum_{kl:\ neighbors\ of\ ij} X_{kl} \\ &= X_{i,j+1} + X_{i,j-1} + X_{i+1,j} + X_{i-1,j} \end{aligned}$$

and $J \in \mathbb{R}$ is a scalar parameter. We already have the likelihood, that is,

$$p(Y_{ij}|X_{ij}) = \mathcal{N}(Y_{ij}; X_{ij}, \sigma)$$

1. Use Bayes' theorem to obtain the full conditional distribution $p(X_{ij}|Y_{ij}, X_{-ij})$. Write the expressions of full conditional probabilities (using the Gaussian likelihood and prior) $p(X_{ij} = 1|Y_{ij}, X_{-ij})$ and $p(X_{ij} = -1|Y_{ij}, X_{-ij})$.
2. Set $\sigma = 4$ to obtain the noisy image in Fig. 1.
3. Implement the Gibbs sampler and sample from the full conditionals. At the end, you will have a denoised image. As a parameter of the prior, $J = 4$ gives a good result. Also try with $J = 0$ and $J = -4$ and report these results. Comment on the results and effect of the parameter J.
4. Test both images given in the webpage: `http://www.cmpe.boun.edu.tr/~akyildiz/teaching/`.

2.3 Life saving details

- Use `imagesc` to visualise the image since we are working on unusual image values such as $[-1, 1]$. Never use functions for images such as `imshow, im2double` etc. To visualise the image Y, use `imagesc(Y)`.

- Create your noisy image with the following script,

```
clc;
clear all;
close all;

X = imread('carriage-17.gif');

ind1 = find(X > 0);
ind2 = find(X == 0);

X = double(X);

X(ind1) = 1;
X(ind2) = -1;

sig = 4;

Y = X + sqrt(sig) * randn(size(X));
```

  where X is the original image, Y is the observed, noisy image.

- To process the rice image, you have to convert the image from RGB to Binary. When you read image with `imread`, you will see that image's dimension is something like $m \times n \times 3$. You can convert the image to the grayscale via `rgb2gray` function. Use the following code to convert image to gray scale.

```
clc;
clear all;
close all;

X = imread('riceImage.png');

X = rgb2gray(X);

ind1 = find(X > 0);
ind2 = find(X == 0);

X = double(X);

X(ind1) = 1;
X(ind2) = -1;

sig = 4;

Y = X + sqrt(sig) * randn(size(X));
```

- A common practical issue in image processing occurs when processing boundaries. To make operations on boundaries, you need extra pixels. If you assume 'outside' of the image is zero, this causes problems. Instead, you can 'extend' the image (it is called padding). You can reflect the boundaries (copying them) and create new pixels. You can either do this by hand or by using the `padarray` function (use at your own risk).

3 Three simple Gibbs samplers

As Taylan hoca said in the lecture, derive Gibbs samplers for three toy models: (1) Toy example in the Gibbs sampler slides (sum of two random variables), (2) Toy example covered in the lecture (Poisson model), (3) Your own toy probability model.