CS 294-2 Grouping and Recognition — 9/13/99
Lecture 7 (Gibbs Distribution, MRF, MCMC) — Scribes Notes by Vito Dai

Gibbs Distribution

$$\pi(\omega) = \frac{1}{z} e^{\frac{-U(\omega)}{T}}$$

Proving Gibbs Distribution Implies Markov Random Field

1. $P(X = \omega) > 0$

   Trivial because of exponential

2. $P(X_s = x_s | X_r = x_r, r \neq s) = P(X_s = x_s | X_r = x_r, r \in G_s)$

$$P(X_s = x_s | X_r = x_r, r \neq s)$$

$$= \frac{P(X_s = x_s, X_r = x_r, r \neq s)}{P(X_r = x_r, r \neq s)}$$
$$\vdots$$
$$= \frac{e^{-U(\omega)/T}}{\sum e^{-U(\omega^*)/T}}$$

Look at the handout given in class for a detailed proof.
In summary, at the end of the day, sites without s cancel.

Back to the Image Segmentation Problem

Our image model

true image f $\xrightarrow{+\eta}$ observed image g

line process *l* not observed

Goal

- Given observed image g, find a probability distribution of true image *f*, and the line process *l*.

The *line process estimate* solves the *image segmentation* problem.
The *true image estimate* solves the *image restoration* problem.
Both problems are simultaneously solved!

- Note: only works for piecewise smoothe images $\Rightarrow$ no textures

Our model assumes:

$$P(F = f, L = l) = \frac{e^{-U(f,l)/T}}{Z}$$

$$g = f + \eta$$

Our solution:

$$P(X = \omega | G = g)$$

$$= \frac{P(X = \omega, G = g)}{P(G = g)}$$

prior distribution

$$= \frac{P(G = g | X = w) P(X = \omega)}{P(G = g)}$$

normalizing constant

Interesting term:

$$P(G = g | X = \omega)$$

$$= P(n = f - g)$$

$$= \frac{1}{\left(\sqrt{2\pi}\sigma\right)^n} \exp\left(\frac{-\|\mu - (g - f)\|^2}{2\sigma^2}\right)$$

Where we assume every pixel has independent noise $\eta \sim N(\mu, \sigma)$
e.g. Poisson process noise in CCDs

Result:

$$P\left(X = \omega \middle| G = g\right)$$

$$= \frac{e^{-U(f,l) + term/T}}{Z'}$$

$$= \frac{e^{-U_p(f,l)/T}}{Z'}$$

is the posterior probability of a particular f, l given g. Note this is also a Gibbs distribution!

MAP (maximum a posteriori) Estimate

If you insist on a single answer then return f*, l* that maximizes

$$P\left(X = \omega^* \middle| G = g\right) = \frac{1}{Z}\exp\left(\frac{-U_p\left(f^*, l^*\right)/T}{Z'}\right)$$

or equivalently, minimizes the *energy function*

$$U_p\left(f^*, l^*\right)$$

Problem: *f* , *l* space is very large!!
Solution: Construct samples of *f* , *l* in this space with high probablity
Technique: Markov Chain Monte Carlo (MCMC) lets you sample the posterior distribution

Sampling a Distribution

Q: How do we represent a probability distribution with sampling?
A: Create many samples drawn from that distribution and count!

Example:

Q: $P(X > 17) = ?$
A: Create samples X_i drawn from the distribution of X.

$$P\left(X > 17\right) = \frac{N\left(x_i > 17\right)}{N_{total}}$$

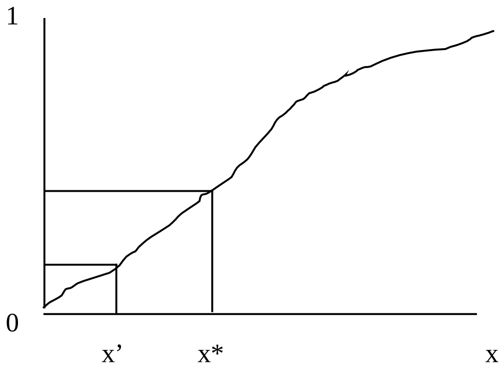
Count the number of samples greater than 17 and divide by total number of samples.

Generating the Samples

Primitive random number generator X ~ U(0,1).
To create Y ~ U(a,b) use

Y = a + (b – a) X

In general we can use the cumulative distribution function



1987 – Stochastic Simulation (Ripley) for generating samples for "standard stuff" in textbooks

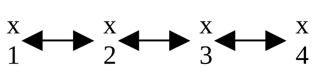
Markov Chain Monte Carlo (MCMC) Technique

- Define a suitable Markov Chain whose equilibrium distribution is the desired posterior distribution
- Generate samples from the Markov Chain

Markov Chain Basics

$$P\left(X_t = y \middle| X_{t-1}, X_{t-2}, \cdots, X_0\right) = P\left(X_t = y \middle| X_{t-1}\right)$$

Example:



$$P\left(X_t = 3 \middle| X_{t-1} = 2\right) = \tfrac{1}{2}$$
$$P\left(X_t = 1 \middle| X_{t-1} = 2\right) = \tfrac{1}{2}$$
$$P\left(X_t = 4 \middle| X_{t-1} = 2\right) = 0$$
$$P\left(X_t = 2 \middle| X_{t-1} = 2\right) = 0$$

This transition probabilities can be written as a matrix

$$P = \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix} \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 1/2 & 1/2 \end{bmatrix}$$

If we write the probability distribution at time t as p(t) then

p(t+1) = p(t)P

For example if the drunk's walk starts at position 2 we denote
p(t = 0) = [0 1 0 0]

p(t) is an evolving probability distribution which is a row vector that sums to one

The equilibrium distribution p(infinity) = π satisfies

$\pi P = \pi$

and is a left eigenvector of P with eigenvalue one.

Finding the Markov Chain Corresponding to the Posterior Distribution

Metropolis Sampler – Rosenberg, Teller, Teller
Heat Bath (Gibbs Sampler) – "rapidly mixing" determines convergence rate

Metropolis Sampler

We are given that the posterior distribution is of the form f(x)/Z

1. We have a proposal kernal satisfying K(x,y) = K(y,x)
2. Calculate f(y)
3. Accept transition with probability = min {1, f(y)/f(x) }
4. This eventually converges to the "right thing"