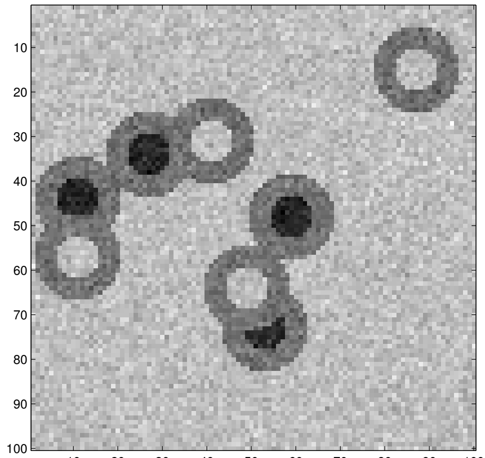
Blockkurs on Markov Chain Monte Carlo for Inverse Problems in PDEs
Computing 4

The following picture shows a (synthetic) noisy scene containing 'good' (black inside) and 'bad' (white inside) cells. Your task is to write an MCMC that automatically counts the number of good and bad cells, and displays the marginal posterior distribution over cell counts.



The image is a 100×100 pixel image, scaled between 0 (black) and 1 (white), in the file `slide.tif` on the Blockkurs website – though the `.tif` file is clipped and scaled to `UINT8`. Also there is the code (`makefake.m`) that produced the original image.

Prior distribution: The 'true' image is defined by an ordered set of $r = (x, y)$ points (centers of cells) (I have used pixel count as the units), with each point being marked with a label ('good' or 'bad'). You should model the *number* of cells as unknown, so random. A sensible choice is a Poisson distribution over number, with some sensible mean (you can investigate this hyperparameter later). The *location* of each cell can be modelled as uniformly distributed within the image, and the label is equally probable to be 'good' or 'bad'. Write down the PDF for the prior distribution that this model implies.

Likelihood function: By looking at the added noise term (in `makefake.m`), write down the likelihood function.

Coding steps: You can break the coding tasks into:

- Start with the template for MCMC with multiple moves in `counting_MCMC_template.m`
- Decide on a data structure for the state X. I suggest $X = \{n, (r_1, m_1), (r_2, m_2), \ldots, (r_n, m_n)\}$ with order setting back to front in the image.
- Write some code to take state X and produce the associated image, using `putgood.m` and `putbad.m` to put good and bad cells in the image..
- Evaluate the log likelihood (from your simulated data), i.e., evaluate $\pi(y|X)$.
- Calculate the log prior for X (it's just the Poisson on n), $\pi(X)$.
- Write proposal moves for your MCMC – each function needs to return the proposed state, and the log of the Hastings ratio (ratio of reverse/forward proposal densities).

  1. A birth-death move that creates or deletes cells.
  2. A swap move, that flips the label of a cell from 'good' to 'bad', and vice versa.
  3. A translate move, that randomly translates the centre of a cell within a window.
  4. A permute move that swaps the order (front to back) of a pair of cells.
  5. Anything else you can think of that seems a good idea.

- Run your MCMC and have fun!

Plot the marginal posterior distribution over the number of good and bad cells.

Challenge question – for the brave: Put a (hyper) prior distribution over the mean number of cells, and present results that integrate over this nuisance parameter. Does this make the counting more robust? (Simulate more data sets to test this in a frequentist sense.)

Colin Fox
26 Jul 2019