Decrypting ciphers using Markov Chain Monte Carlo

Course Project for Statistical Computing

Cosmin Borsa[*]*, Yichu Li*[†]

10/12/2018

Abstract

In this project we are going to implement a Markov Chain Monte Carlo algorithm to decode a classical cipher text.

We are going to use the work of Jian Chen, Jeffrey S. Rosenthal, Andrew Landgraf and Persi Diaconis to implement a Markov Chain Monte Carlo algorithm that decrypts coded messages. We are going to decode are simple substitution codes, in which each symbol of the coded message stands for a letter of the English alphabet. We will attempt to decode the cipher by finding a decoding function f whose output generates the plain text.

$$f : \{\text{Cipher text}\} \mapsto \{\text{Plain text}\}$$

To solve this problem, we will use the statistics of written English, in particular the frequency analysis of pairs of letters to crack the coded message. To do so we will use a reference text such as the novel *War and Peace* by Leo Tolstoy to construct a matrix $M(a, b)$ with the transition frequencies from each letter to the next.

To find the best decoding function f for the cipher we will use a plausibility function Pl. The plausibility function Pl assess the goodness of f by assigning to a sequence of coded symbols in the cipher the frequency with which the consecutive pairs of decrypted symbols appear in the reference text. High values of $Pl(f)$ will indicate that the decoding function f is good in assigning pairs of letters while low values point to the fact that f is assigning consecutive symbols in the cipher to pairs of letters which are unlikely to occur in standard written English.

$$Pl(f(s_1 s_2 s_3 \ldots)) = \prod_i M(f(s_i), f(s_{i+1}))$$

We will find a decoding function f from an initial guess and by iterating a Metropolis algorithm many times (e.g. $2,000$). At first, the decoding function f is used to compute the value of the plausibility function for the sequence of symbols. Second, a random transposition is made to the decoding function f; for example, if f mapped the symbol $s_i \mapsto A$ and $s_{i+1} \mapsto H$, the new decoding function f^* will map $s_{i+1} \mapsto A$ and $s_i \mapsto H$. Next, the value of $Pl(f^*)$ is calculated for the same sequence of symbols. If $Pl(f^*)$ is higher than $Pl(f)$, then function f^* will be used as a decoding function, replacing f. On the other hand, if $Pl(f^*)$ is lower than the value of the plausibility function of f, a Bernoulli trial is conducted. If the experiment generates a success, the function f^* will be used as a decoding function and it will take the place of the function f; otherwise, the

[*]cosmin.borsa@uconn.edu; M.S. in Applied Financial Mathematics, Department of Mathematics, University of Connecticut.

[†]yichu.li@uconn.edu; M.S. in Applied Financial Mathematics, Department of Mathematics, University of Connecticut.

function f^* will be discarded. The Bernoulli trial is run since the decoding function might hit a local maximum. In order to avoid getting stuck at a local maximum, it is reasonable to allow the decoding function to temporally take lower values such that in subsequent iterations the function may reach or come close to the actual maximum. After a decision regarding f^* has been made, the process is repeated for another random transposition of the decoding function. Finally, after $2,000$ or more iterations, the algorithm should deliver a good decoding function f with which the message might be decrypted.

References

[1] Persi Diaconis. *The Markov Chain Monte Carlo Revolution*. Bulletin of the American Mathematical Society, Volume 6, Number 2, Pages 179 - 205, 2009.

[2] Jian Chen, Jeffrey S. Rosenthal. *Decrypting classical cipher text using Markov chain Monte Carlo*. Statistics and Computing, Volume 22, Issue 2, Pages 397 - 413, 2012.

[3] Andrew Landgraf: Text Decryption Using MCMC, `http://alandgraf.blogspot.com/2013/01/text-decryption-using-mcmc.html`