Fast Monte Carlo Method for Condition Number Estimation

Silvi-Maria Gurova[0000-0002-5026-6910], Aneta Karaivanova

Institute of Information and Communication Technologies, Bulgaria Academy of Sciences,
Acad. G. Bonchev St., bl.25A, Sofia, Bulgaria
smgurova@parallel.bas.bg, anet@parallel.bas.bg

Abstract. This paper presents theoretical results on estimation the condition number with application in image restoration problems.

Keywords: Monte Carlo Method, Markov Chain, Condition Number, Eigenvalue.

1 Introduction

The advances in algorithms, visualization, data analytics are expected to greatly impact progress in the digital presentation and preservation of cultural heritage which is of enormous importance for the human society. This paper presents theoretical results on estimating the condition numbers which is an important problem in image restoration problems. We follow the problem description in (Stanley, Ankit, Truong, & Edmund, 2011):

Image restoration is an inverse problem where the goal is to recover a sharp image from a blurry and noisy observation. Using the classical shift-invariant imaging system model (Kim, 2002), the input-output relationship is given by:

$$g = Hf + h, \qquad (1)$$

where f is a vector denoting the unknown (potentially sharp) image, g is the observed blurry and noisy image, h is the noise vector and H is a matrix that models the blur (convolution matrix). If the blur is spatially invariant, meaning that all pixels of the image f are identically blurred, then the matrix H has a block-circulant-with-circulant-block (BCCB) structure (Chan, 2010). In this case, H can be diagonalized using Fourier transforms. However, if the blur is *spatially variant*, then H is *not diagonalizable* using DFT matrices, thus making the spectral analysis of H difficult. In (Stanley, Ankit, Truong, & Edmund, 2011) a method for estimation the upper and lower bounds on the largest and smallest eigenvalues of $H^T H$ and hence the condition number of $H^T H$ is presented.

The goal of this paper is to present another method, namely a Fast Monte Carlo method for estimation the condition number of a given matrix.

Digital Presentation and Preservation of Cultural and Scientific Heritage. Conference Proceedings. Vol. 11, Sofia, Bulgaria: Institute of Mathematics and Informatics – BAS, 2021. ISSN: 1314-4006, eISSN: 2535-0366

2 Background

2.1 Formulation

The condition number of square nonsingular matrix *A* is defined by:

$$cond(A)=||A||.||A^{-1}||. \tag{1}$$

The numerical value of the condition number of an nxn matrix depends on the particular norm used but because of the equivalence of the underlying vector norms, these values can differ by at most a fixed constant and hence they are equally useful as quantitative measure of conditioning number.

In this paper we use the following matrix norm

$$||A||_2 = \sigma_{max}(A), \tag{2}$$

where $\sigma_{max}(A)$ represents the largest singular value of matrix *A*. A singular value of a real matrix *A* is the positive square root of an eigenvalue of the symmetric matrix AA^T or A^TA. For symmetric and Hermitian matrices the eigenvalues and singular values are obviously closely related. A nonnegative eigenvalue, $\lambda \geq 0$, is also a singular value where $\sigma(A) = \lambda$. Next, if a matrix *A* is diagonalizable, then the eigenvalues of the inverse matrix are reciprocals of the eigenvalues of *A*.

Let $A = \{a_{ij}\}_{i,j=1}^{n} \in \mathbb{R}^{nxn}$ be a given (non-singular) matrix. The eigenvalue problem which we consider is to find $\lambda(A)$ such that:

$$Au = \lambda(A)u. \tag{3}$$

A matrix polynomial of degree *k* is donated by the equation:

$$p_k(A) = \sum_{i=0}^{k} c_i A^i, \qquad c_i \in \mathbb{R}. \tag{4}$$

In the rest of the paper we concentrate on computing the largest and the smallest eigenvalues of a matrix using Monte Carlo (MC) method.

2.2 Computing the Extreme Eigenvalues

The well-known Power method (Golub & Van Loon, 1996), (Isaacson & Keller, 1996) gives an estimate for the dominant eigenvalue λ_1. Let *A* be an nxn matrix with real elements a_{ij}. We denote the matrix eigenvectors by $\{u_j\}_{j=1}^{n}$ and suppose that

$$|\lambda_1| > |\lambda_2| \geq \ldots\ldots \geq |\lambda_{n-1}| > |\lambda_n|.$$

Using the Power method for estimating the eigenvalue, we get

$$\lambda_{max}^{(k+1)} = \lambda_1 \frac{a_1 u_{i1} + \sum_{j=2}^{n} (\lambda_j / \lambda_1)^{k+1} a_j u_{ij}}{a_1 u_{i1} + \sum_{j=2}^{n} (\lambda_j / \lambda_1)^{k} a_j u_{ij}}, \tag{5}$$

$$\lambda_{max}^{(k+1)} = \lambda_1 + O(|\lambda_2 / \lambda_1|^k) \tag{6}$$

If *A* is a symmetric matrix and has real eigenvalues and eigenvectors, formulas (5) and (6) become:

$$\mu_{k+1} = \lambda_1 \frac{|a_1|^2 + \sum_{j=2}^{n} |a_j|^2 (\lambda_j / \lambda_1)^{2k+1}}{|a_1|^2 + \sum_{j=2}^{n} |a_j|^2 (\lambda_j / \lambda_1)^{2k}}, \tag{7}$$

$$\mu_{k+1} = \lambda_1 + O(|{}^{\lambda_2}/_{\lambda_1}|^{2k}) \quad (8)$$

There are two deterministic numerical methods that can efficiently compute only the extremal eigenvalues: *the Power method* and *the Lanczos-type method*. We should note that the Lanczos-type method is used only for symmetric eigenvalue problems (Golub & Van Loon, 1996).

Computational Complexity: If *k* iterations are required for convergence, the number of arithmetic operations is $O(kn^2)$ for the Power method and $O(n^3 + kn^2)$ for both methods – the Inverse and Inverse Shifted Power method.

3 Condition Number Estimation Using Monte Carlo

3.1 Monte Carlo Method for Approximate Computing of Eigenvalues

Let $h = \{h_i\}_{i=1}^n$ and $f = \{f_i\}_{i=1}^n$ be two arbitrary vectors in $\mathbb{R}^n$. To create a stochastic process we use the matrix *A* and those two given vectors. The problem we are interested in evaluating is the scalar products of the following type:

$$(h, p_m(A)f). \quad (9)$$

Considering the special case $p_m(A) = A^m$, the form (9) becomes

$$(h, A^m f). \quad (10)$$

The problem of approximate computing of the largest eigenvalue can be reformulated as:

$$\lambda_1 \approx \lim_{m\to\infty} \frac{(h,A^m f)}{(h,A^{m-1}f)}.$$

However, we are also interested in finding the smallest eigenvalue of the matrix *A*, so we work with its *resolvent* matrix $R_q = [I - qA]^{-1} \in \mathbb{R}^{n\times n}$ (Dimov & Karaivanova, 1998). If $|q\lambda| < 1$, R_q may be expanded as a series via binomial theorem:

$$p_\infty = p(A) = \sum_{i=0}^{\infty} q^i C_{i+m-1}^i A^i = [I - qA]^{-m} = R_q^m. \quad (11)$$

A connection between the eigenvalues of the matrices R_q and *A* existed and its given by the equality:

$$\mu = \frac{1}{1-q\lambda} \quad (12)$$

and the eigenvectors of the two matrices are coincided. If $q>0$, the largest eigenvalue μ_{max} of the *resolvent* matrix corresponds to the largest eigenvalue λ_{max} of the matrix *A*. But when $q<0$, then μ_{max} corresponds to the smallest eigenvalue λ_{min} of the matrix *A*.

Applying the Power method (Dimov & Karaivanova, 1998), leads to the following iterative processes:

$$\lambda^{(m)} = \frac{(h,A^m f)}{(h,A^{m-1}f)} \xrightarrow[m\to\infty]{} \lambda_{max}\ , \quad (13)$$

$$\mu^{(m)} = \frac{([\mathrm{I}-\mathrm{qA}]^{-\mathrm{m}}f,h)}{([\mathrm{I}-\mathrm{qA}]^{-(\mathrm{m}-1)}f,h)} \xrightarrow[m\to\infty]{} \mu_{max} = \frac{1}{1-q\lambda} \quad (14)$$

To construct the Monte Carlo (MC) method, we define a Markov chain

$$k_0 \to k_1 \to \cdots k_i \to \cdots, \quad (1 \le k_i \le n)$$

with initial density vector $p = \{p_\alpha\}_{\alpha=1}^{n}$, where $Pr(k_i = \alpha) = p_\alpha$ and the transition density matrix $P=\{p_{\alpha\beta}\}_{\alpha,\beta=1}^{n}$, where $Pr(k_j = \beta | k_{j-1} = \alpha) = p_{\alpha\beta}$.

Define the following random variables (r.v.) W_j using the recursion formula:

$$W_o = \frac{h_{k0}}{p_{k0}}, \quad W_j = W_{j-1}\frac{a_{k_{j-1}k_j}}{p_{k_{j-1}k_j}}, \qquad j = 1,2,\ldots \tag{15}$$

This has the desired expected values:

$$E\left[W_i f_{k_i}\right] = (h, A^i f), \quad i = 1,2\ldots. \tag{16}$$

$$E\left[\sum_{i=0}^{\infty} q^i C_{i+m-1}^{i} W_i f(x_i)\right] = (h, [I - qA]^{-m} f), \quad m = 1,2\ldots. \tag{17}$$

This allows us to estimate the desired eigenvalues as:

$$\lambda_{max} \approx \frac{E\left[W_i f_{k_i}\right]}{E\left[W_{i-1} f_{k_{i-1}}\right]} \tag{18}$$

and

$$\lambda \approx \frac{1}{q}\left(1 - \frac{1}{\mu^{(m)}}\right) = \frac{E[\sum_{i=1}^{\infty} q^{i-1} C_{i+m-2}^{i-1} W_i f(x_i)]}{E[\sum_{i=0}^{\infty} q^i C_{i+m-1}^{i} W_i f(x_i)]}. \tag{19}$$

In formula (18) the length of the Markov chain is equal to the number of iterations, *i,* in the Power method while in formula (19) the length of the Markov chain is equal to the number of terms in truncated series for the r*esolvent* matrix, where the parameter *m* corresponds to the number of power iterations.

3.2 Estimation of the Condition Number

By the equality (12), we can calculate the approximate values of λ_{max} and λ_{min}, using the already approximated extremal eigenvalues of the *resolven*t matrix

$$\lambda \approx \frac{1}{q}\left(1 - \frac{1}{\mu^{(m)}}\right).$$

For the Resolvent Power method in case when $q < 0$ the eigenvalues of matrices A and R_q are connected through the equality:

$$\lambda_{n-i+1} = \frac{1}{|q|}\left(\frac{1}{\mu_i} - 1\right) \tag{20}$$

Then, $\lambda_{min} = \lambda_n = \frac{1}{|q|}\left(\frac{1}{\mu_1} - 1\right)$ for $i = 1$ and `$\lambda_{max} = \lambda_1 = \frac{1}{|q|}\left(\frac{1}{\mu_n} - 1\right)$ for $i = n$, where $\mu_1 = \mu_{max}$ and $\mu_n = \mu_{min}$. Thus, we can estimate the condition number of the considered matrix as:

$$cond(A) = \frac{|\lambda_{max}|}{|\lambda_{min}|} = \frac{|\mu_{max}|}{|\mu_{min}|}\frac{|1-\mu_{min}|}{|1-\mu_{max}|} = cond(R_q)\frac{|1-\mu_{min}|}{|1-\mu_{max}|} \tag{21}$$

The last formula (21) shows that $cond(A) < cond(R_q)$ when $\frac{|1-\mu_{min}|}{|1-\mu_{max}|} < 1$ and vice versa. If we have this apriori estimation, we can choose which one of the two methods, Power Monte Carlo or Resolvent Power Monte Carlo method, to use in our computation in order to solve faster the eigenvalue problem.

3.3 Convergence and Complexity

Consider a random variable $\theta^i = W_i f_{ki}$ that has a mathematical expectation formula (16). The MC error obtained when computing a matrix-vector product is well-known to be:

$$\left|h^T A^i f - \frac{1}{N}\sum_{i=0}^{N}(\theta^i)_s\right| \approx c.Var(\theta^i)^{1/2}N^{-1/2},$$

where c is a constant, the $\mathrm{Var}(\theta^i) = \mathrm{E}[(\theta^i)^2]-(\mathrm{E}[\theta])^2$ and

$E[\theta^i] = \mathrm{E}[\frac{h_{k0}}{p_{k0}} W_i f_{k_i}] = \sum_{k_0=1}^{n} \frac{h_{k0}}{p_{k0}} p_{k_0} \sum_{k_1=1}^{n} \dots \sum_{k_i=1}^{n} \frac{a_{k_0k_1\dots a_{k_{m-1}k_m}}}{p_{k_0k_1\dots p_{k_{m-1}k_m}}} p_{k_0k_1} \dots \cdot p_{k_{m-1}k_m}.$

Using the Importance sampling, we define the initial and transition densities as follows:

$$p_\alpha = \frac{|h_\alpha|}{\sum_{\alpha=1}^{n}|h_\alpha|} \text{ and } p_{\alpha\beta} = \frac{|h_{\alpha\beta}|}{\sum_{\beta=1}^{n}|h_{\alpha\beta}|}, \quad \alpha = \overline{1,n}.$$

In this case we get the following estimation for the variance:

$$Var[\theta^i] = E\left[\left(h_{k_0}W_m f_{k_m}\right)^2\right] - (E[h_{k_0}W_m f_{k_m}])^2 \le E\left[\left(h_{k_0}W_m f_{k_m}\right)^2\right] \le \sum_{i=1}^{n}|a_{k_0 i}| \sum_{i=1}^{n}|a_{k_1 i}| \dots \sum_{i=1}^{n}|a_{k_{m-1} i}|,$$

for any f and h that are normalized.

The special case is when the elements of A are positive and the rows' sums of A are constant, i.e $\sum_{j=1}^{n} a_{ij} = a$, $i=\overline{1,n}$ and if all the elements of the vector f are constant, then $Var[\theta^i] = 0$ (Dimov, Philippe, Karaivanova, & Weihrauch, 2008).

The convergence of the Resolvent MC method is $O\left(\left|\mu_1/\mu_2\right|^m + cN^{-\frac{1}{2}}\right)$, where m corresponds to the number of power iterations with the *resolvent* matrix. The computational complexity of the Resolvent MC method is $4lN$, where N is the number of Markov chains, and l is equal to the number of terms in truncated series for the resolvent matrix.

4 Conclusion

Why we are interested in MCMs for eigenvalue problem? Because the computational complexity is bounded by *O(lN)* where N is the number of chains, and l is the mathematical expectation of the length of the chains, both of which are independent of the matrix size n. This makes MCMs very efficient for large, sparse problems for which deterministic methods are not computationally efficient. Our future plans include numerical estimating the condition numbers in image restoration problems.

Acknowledgements

The work was supported by the European Commission through the H2020 Research Infrastructures under the project NI4OS -Europe -National initiatives for open science in Europe (H2020 № 857645) and also has been accomplished with the partial support by the Grant No. BG05M2OP001-1.001-0003, financed by the Science and Education for Smart Growth Operational Program (2014-2020) and co-financed by the European Union through the European structural and Investment funds.

References

Chan, S. (2010). Constructing a sparse convolution matrix for shift varying image restoration problems. *ICIP 2010.* in Proc. IEEE.

Dimov, I., & Karaivanova, A. (1998). Parallel computations of eigenvalues based on a Monte Carlo approach. *Monte Carlo Methods and Applications, 4*(1), 33-52.

Dimov, I., Philippe, B., Karaivanova, A., & Weihrauch, C. (2008). Robustness and applicability of Markov chain Monte Carlo algorithms for eigenvalue problems. *Applied Mathematical Modelling, 32*(8), 1511-1529. doi:http://dx.doi.org/10.1016/j.apm.2007.04.012

Golub, G., & Van Loon, C. (1996). *Matrix Computation.* Baltimore: The Johns Hopkins University Press.

Isaacson, E., & Keller, H. (1996). *Analysis of Numerical Methods.* Dover Publications, INC.N.Y.

Kaposi, G., Szkaliczki, T., Márkus, Z. L., Luchev, D., Goynov, M., & Paneva-Marinova, D. (2013). Mobile Exploring of the Bulgarian Iconography through QR Codes in the GUIDE@HAND Tourist Guide Application. *International Conference on Digital Presentation and Preservation of Cultural and Scientific Heritage, September 18-21, 2013, Veliko Tarnovo, Bulgaria. III*, pp. 44-52. Sofia: Institute of Mathematics and Informatics - BAS.

Kim, B. (2002). Numerical Optimization Methods for Image Restoration. Stanford University.

Stanley, H. C., Ankit, K., Truong, Q. N., & Edmund, Y. L. (2011). Bounds for the Condition Numbers of Spatially-variant Convolution Matrices in Image Restoration Problems. *SMA4.* Optical Society of America. doi:https://doi.org/10.1364/SRS.2011.SMA4

Received: June 25, 2021
Reviewed: July 18, 2021
Finally Accepted: July 27, 2021