A Randomized Algorithm for Approximating the Log Determinant of a Symmetric Positive Definite Matrix

Eugenia-Maria Kontopoulou

in collaboration with

C. Boutsidis, P. Drineas, P. Kambadur & A. Zouzias

Department of Computer Science
Purdue University

CSESC 2017 Conference

Purdue, April 2017

The problem of $\log\det(\mathbf{A})$

Definition

Given a Symmetric Positive Definite matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$, compute (exactly or approximately) the $\log\det(\mathbf{A})$.

Application: Maximum likelihood estimations, Gaussian processes prediction, log det-divergence metric, barrier functions in interior point methods . . .

Straightforward Computation

1. Compute the Cholesky Factorization of $\mathbf{A}$, and let $\mathbf{L}$ be the Cholesky factor.
2. Compute the log-determinant of $\mathbf{A}$ using $\mathbf{L}$:

$$\log\det(\mathbf{A}) = \log\det(\mathbf{L})^2 = 2\log\prod_{i=1}^{n} L_{ii}.$$

Time Complexity: $\mathcal{O}(n^3)$.

Prohibitive for Large Data!!!!

Lemma

Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ *be an SPD matrix. For any* α *with* $\lambda_1(\mathbf{A}) < \alpha$*, define* $\mathbf{B} = \mathbf{A}/\alpha$ *and* $\mathbf{C} = \mathbf{I}_n - B$*. Then,*

$$\log\det(\mathbf{A}) = \underbrace{n\log(\alpha)}_{\Delta_1} - \underbrace{\sum_{k=1}^{\infty} \frac{\mathbf{Tr}\left(\log\left(\mathbf{C}^k\right)\right)}{k}}_{\Delta_2}.$$

Algorithm 1

Input: $\mathbf{A} \in \mathbb{R}^{n \times n}$, accuracy parameter $\epsilon > 0$, integer $m > 0$.

Output: $\widehat{\log\det}(\mathbf{A})$, the approximation to the $\log\det(\mathbf{A})$.

1. Compute an estimate to the largest eigenvalue of $\mathbf{A}$, $\tilde{\lambda}_1(\mathbf{A})$, using the Power Method.
2. $\mathbf{C} = \mathbf{I}_n - \mathbf{A}/(7\tilde{\lambda}_1(\mathbf{A}))$
3. Create $p = \lceil 20\log(2/\delta)/\epsilon^2 \rceil$ i.i.d random Gaussian vectors, $\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_p$.
4. Estimate Δ_2 with a truncated Taylor Series type randomized trace estimator that computes:

$$\Delta_2 \approx \sum_{k=1}^{m} \left(\frac{1}{p} \sum_{i=1}^{p} \mathbf{g}_i^\top \mathbf{C}^k \mathbf{g}_i \right).$$

Lemma

Let $\widehat{\log\det}(\mathbf{A})$ *be the* $\log\det$ *approximation of the above procedure. Then, we* **prove** *that with probability at least* $1-2\delta$*,*

$$\left|\widehat{\log\det}(\mathbf{A}) - \log\det(\mathbf{A})\right| \le 2\epsilon\Gamma$$

where $\Gamma = \sum_{i=1}^{n} \log\left(7 \cdot \frac{\lambda_1(\mathbf{A})}{\lambda_i(\mathbf{A})}\right)$ *and* $m \ge \lceil 7\kappa(\mathbf{A}) \log(\frac{1}{\epsilon}) \rceil$*.*

Running Time

$$\mathcal{O}\left(nnz(\mathbf{A}) \cdot \left(\frac{m}{\epsilon^2} + \log n\right) \cdot \log\left(\frac{1}{\delta}\right)\right).$$

Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be an SPD matrix whose eigenvalues lie in the interval $(\theta_1, 1)$, for some $0 < \theta_1 < 1$. Let $\mathbf{C} = \mathbf{I}_n - A$. Then,

$$\log\det(A) = -\underbrace{\sum_{k=1}^{\infty} \frac{\mathbf{Tr}\left(\log\left(\mathbf{C}^k\right)\right)}{k}}_{\Delta}.$$

Algorithm 2

Input: $\mathbf{A} \in \mathbb{R}^{n \times n}$, accuracy parameter $\epsilon > 0$, integer $m > 0$.

Output: $\widehat{\log\det}(\mathbf{A})$, the approximation to the $\log\det(\mathbf{A})$.

1. $\mathbf{C} = \mathbf{I}_n - \mathbf{A}$
2. Create $p = \lceil 20 \log(2/\delta)/\epsilon^2 \rceil$ i.i.d random Gaussian vectors, $\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_p$.
3. Estimate Δ with a truncated Taylor Series type randomized trace estimator as:

$$\Delta \approx \sum_{k=1}^{m} \left(\frac{1}{p} \sum_{i=1}^{p} \mathbf{g}_i^\top \mathbf{C}^k \mathbf{g}_i \right).$$

Lemma

Let $\widehat{\log\det}(\mathbf{A})$ *be the* $\log\det$ *approximation of the above procedure on inputs* $\mathbf{A}$ *and* ϵ*. Then, we* **prove** *that with probability at least* $1-\delta$*,*

$$\left|\widehat{\log\det}(\mathbf{A}) - \log\det(A)\right| \leq 2\epsilon \cdot \left|\log\det(\mathbf{A})\right|$$

and $m \geq \lceil \frac{1}{\theta_1} \cdot \log(\frac{1}{\epsilon}) \rceil$.

Running Time

$$\mathcal{O}\left(\frac{\log(1/\epsilon)\log(1/\delta)}{\epsilon^2\theta_1} \cdot nnz(\mathbf{A})\right).$$

n	$\log\det(\mathbf{A})$			**time (secs)**		
	exact	**mean**	**std**	**exact**	**mean**	**std**
5000	-3717.89	-3546.920	8.10	2.56	1.15	0.0005
7500	-5474.49	-5225.152	8.73	7.98	2.53	0.0015
10000	-7347.33	-7003.086	7.79	18.07	4.47	0.0006
12500	-9167.47	-8734.956	17.43	34.39	7.00	0.0030
15000	-11100.9	-10575.16	15.09	58.28	10.39	0.0102

Table: Parameters: $p = 60$, $m = 4$, $t = \log(\sqrt{4n})$. Ground truth computed via Cholesky. Mean and standard deviation reported over 10 repetitions.

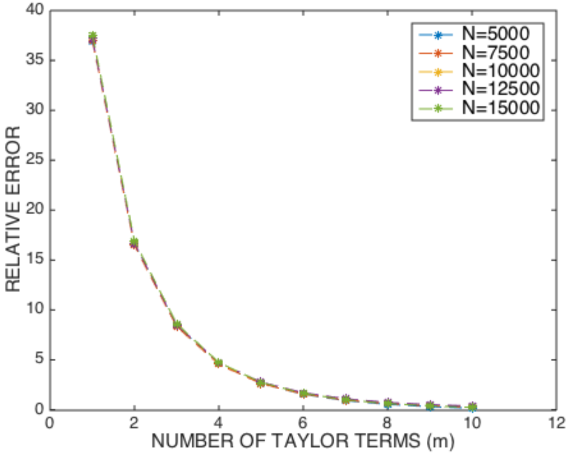

Figure: Parameters: $p = 60$ and $t = 2 \log \sqrt{4n}$. Ground truth computed using Cholesky.

name	n	nnz	log det **A**			**time (sec)**		m
			exact	**approx**		**exact**	**approx**	
				mean	**std**		**mean**	
thermal2	1228045	8580313	1.3869e6	1.3928e6	964.79	31.28	31.24	149
ecology2	999999	4995991	3.3943e6	3.403e6	1212.8	18.5	10.47	125
ldoor	952203	42493817	1.4429e7	1.4445e7	1683.5	117.91	17.60	33
thermomech_TC	102158	711558	-546787	-546829.4	553.12	57.84	2.58	77
boneS01	127224	5516602	1.1093e6	1.106e6	247.14	130.4	8.48	125

Table: Parameters: $p = 5$, $m = 1 : 5 : 150$ and select the one with best avg, $t = 5$). Ground truth computed via Cholesky. Mean reported over 10 repetitions.

Questions?

Bibliography

C. Boutsidis, P. Drineas, P. Kambadur, E. Kontopoulou, A. Zouzias (2016), ``A Randomized Algorithm for Approximating the Log Determinant of a Symmetric Positive Definite Matrix'', submitted to LAA

ArXiv: https://arxiv.org/pdf/1503.00374.pdf