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

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

Von-Neumann Entropy
Given a Density Matrix $R \in \mathbb{R}^{n \times n}$, compute (exactly or approximately) the Von-Neumann entropy, $H(R)$.

A Density Matrix is represented by the statistical mixture of pure states and has the form
$$ R = \sum_{i=1}^{n} p_i \psi_i^{\dagger} \psi_i = \sum_{i=1}^{n} p_i \phi_i^{\dagger} \phi_i \in \mathbb{R}^{n \times n}, $$
where the vectors $\psi_i \in \mathbb{R}^n$ represent the pure states of a system and are pairwise orthogonal and normal, while $p_i$’s correspond to the probability of each state and satisfy $p_1 + \cdots + p_n = 1$.

Application: Information theory, quantum mechanics, …

Approximation via Taylor
Let $A \in \mathbb{R}^{n \times n}$ be an SPD matrix whose eigenvalues lie in the interval $(\theta_1, 1)$, for some $0 < \theta_1 < 1$. Let $C = I - A$. Then (using the Taylor expansion of $\log C$),
$$ \log \det(A) \approx -\sum_{k=1}^{\infty} \frac{\text{Tr}(C^k)}{k}, $$

where $\text{Tr}(x)$ is the trace operator, $\text{Tr}(x) = \sum_{i=1}^{n} x_{ii}$.

Let $R \in \mathbb{R}^{n \times n}$ be an SPD matrix with unit trace, whose eigenvalues lie in the interval $[\theta_1, \theta_2]$, for some $0 < \theta_1 < \theta_2 < 1$. Let $C = I - u^{-1} R$. Then
$$ H(R) \approx \log u^{-1} + \sum_{k=1}^{\infty} \frac{\text{Tr}(RC^k)}{k}. $$

Approximation via Chebyshev
We can approximate $h(x) = x \log x$ in the interval $(0, a]$ by
$$ f_n(x) = \sum_{k=0}^{n} a_k T_k(x), $$

where $T_k(x) = \cos(k \cdot \arccos((2x/a) - 1))$, the Chebyshev polynomials of first kind for $t > 0$ and $a_k = \frac{u}{t^{k+1}}$, $a_0 = \frac{u}{t}$, $a_{k-1} = \frac{u}{t} a_{k-1}$ for $t \geq 2$. Then
$$ N(R) \approx -\text{Tr}(f_n(R)). $$

Gaussian trace estimator
We use Gaussian trace estimators to estimate the trace of powers of $C$. An $(\epsilon, \sigma)$-Gaussian trace estimator for any SPD $A \in \mathbb{R}^{n \times n}$ is the matrix
$$ G = \frac{1}{p} \sum_{i=1}^{p} A^{1/2} g_i, $$

where the $g_i$’s are $p$ independent random vectors whose entries are i.i.d. standard normal variables. [AvronTalebi2011]

Clenshaw’s Algorithm
We use Clenshaw’s algorithm to evaluate Chebyshev polynomials with matrix inputs. Clenshaw’s algorithm is a recursive approach with base cases $h_0(x) = x$, $h_1(x) = 1$ and the recursive step (for $k = m, m-1, \ldots, 0$) which in our case is: $h_k(x) = a_k x + 2 \left( \frac{\epsilon - 1}{\epsilon} \right) h_{k-1}(x) - h_{k-2}(x)$. Then:
$$ f_n(x) = \frac{1}{2} \left( a_0 + h_n(x) - h_0(x) \right). $$

Power Method
We use the Power Method algorithm to estimate the largest eigenvalue of the matrix. We prove the following, building upon [Trevisan2011]:

Let $p_i$ be the output of Power Method with $q = [4.82 \log(1/\delta)]$ restarts of the algorithm and $t = \frac{\log \log(1/\delta)}{4 \log(1/\delta)}$ iterations before each restart. Then, with probability at least $1 - \delta$,
$$ \frac{1}{\sqrt{p}} \leq p \leq \sqrt{p_i}, $$

where $p$ is the desired upper bound.

Algorithms

Algorithm 1 - LogDet via Taylor Approximation
Input: $A \in \mathbb{R}^{n \times n}$ with eigenvalues lie in $(\theta_1, 1)$ where $\theta_1 > 0$, accuracy parameter $p, q > 0$, integer $m > 0$.
Output: $\log \det(A)$, the approximation to the LogDet of $A$.

1. Compute $p_i = \frac{\log(2 \log(1/p) + 3)}{4 \log(1/p)}$ and $q_i = \frac{\log(2 \log(1/p) + 3)}{4 \log(1/p)}$
2. Compute $p = \max(p_i, q_i, \cdots, p_i, q_i)$ and $q = \max(p_i, q_i, \cdots, p_i, q_i)$
3. Compute $H(R) = \log u^{-1} + \sum_{k=1}^{p} \frac{\text{Tr}(RC^k)}{k}$

Algorithm 2 - Entropy via Taylor Approximation
Input: $R \in \mathbb{R}^{n \times n}$, accuracy parameter $\epsilon > 0$, integer $m > 0$.
Output: $H(R)$, the Taylor approximation to the $H(R)$.

1. Compute $\hat{\mu}_k$ the estimation of the largest singular value of $R$, using power method.
2. Set $u = \min(\{1, 6\epsilon\})$
3. Compute $G = I - u^{-1} R$
4. Compute $H(R) = \log u^{-1} + \sum_{k=1}^{\infty} \frac{\text{Tr}(G C^k)}{k}$

Algorithm 3 - Entropy via Chebyshev Approximation
Input: $R \in \mathbb{R}^{n \times n}$, accuracy parameter $\epsilon > 0$, integer $m > 0$.
Output: $H(R)$, the Chebyshev approximation to the $H(R)$.

1. Compute $\hat{\mu}_k$, the estimation of the largest singular value of $R$, using power method.
2. Set $u = \min(\{1, 6\epsilon\})$
3. Compute $G = I - u^{-1} R$
4. Compute $H(R) = \frac{1}{2} \sum_{k=0}^{m} g_i f_n(R) g_i$

Theoretical Results
We prove that:
- $\log \det(A)$ is an $(\epsilon, \delta)$-estimator of $\log \det(A)$ and can be computed in
  $O \left( \frac{\log(1/\epsilon)}{\epsilon^2 \cdot \Theta(1)} \cdot \text{nnz}(A) \right)$
- $H(R)$ is an $(\epsilon, \delta)$-estimator of $H(R)$ and can be computed in
  $O \left( \frac{\log(1/\epsilon) \log(1/\delta)}{\epsilon^2 \cdot \text{nnz}(R)} + \log n \cdot \log(1/\delta) \cdot \text{nnz}(R) \right)$

Future Directions
- Eigenvalue distribution: How much is the relative error affected by the distribution of the eigenvalues when we use the polynomial-based algorithms?
- Low rank Density Matrices: Polynomial-based algorithms do not work. We can use Random Projections to approximate the Von-Neumann Entropy of low rank Density Matrix. Theoretically we get additive $(\epsilon, \delta)$-estimators. Considerations: Fast construction of the random projection matrix, meaningful bounds…

Matrix of size $16,384 \times 16,384$

Citations
- L. Trevisan, Graph Partitioning and Expanders, Handout 7, 2011.