Semidefinite Approximations of the Matrix Logarithm

CHEN Li

IORA, NUS

Reading Seminar
June 5th, 2020

Outline

Preliminaries

- Semidefinite representable

Definition 1

A convex function f is said to have a semidefinite representation of size d if its epigraph $\{(x,t) : f(x) \leq t\}$ can be expressed in the form $\pi(L \cap \mathbf{H}_+^d)$ where π is a linear map, $L \subseteq \mathbf{H}^d$ is a subspace of $d \times d$ Hermitian matrices and $\mathbf{H}_+^d$ is the Hermitian positive semidefinite cone.

- Essentially, it means we can write the epigraph in terms of SDP
- Matrix function

Definition 2

For a function $g : \mathbb{R}_{++} \to \mathbb{R}$, the corresponding matrix function can be defined for any $X \in \mathbf{H}_{++}^d$ by $g(X) = U \operatorname{diag}(g(\lambda_1), ..., g(\lambda_d)) U^*$ where $X = U \operatorname{diag}(\lambda_1, ..., \lambda_d) U^*$ eigendecomposition of X.

Examples

- Convex quadratic functions: $f(x) = x^T Q x + c^T x + d$ with $Q = R^T R$

$$f(x) \leq t \Longleftrightarrow \begin{bmatrix} t - d - c^T x & x^T R^T \\ Rx & I \end{bmatrix} \succeq 0$$

- X^2 is semidefinite representable since $X^2 \preceq T \Longleftrightarrow \begin{bmatrix} I & X \\ X & T \end{bmatrix} \succeq 0$
- $\|X\|_p = \left(\sum_{i=1}^d |\lambda_i(X)|^p\right)^{1/p}$ where $p \geq 1$ is rational, and more in [Ben-Tal and Nemirovski, 2001]
- Note 2×2 LMI can be modeled as second-order cone programming (SOCP)

$$\begin{bmatrix} x & y \\ y & z \end{bmatrix} \succeq 0 \Longleftrightarrow x + z \geq \sqrt{(x-z)^2 + 4y^2}, x \geq 0, z \geq 0$$

Problem

- Question: Is $\log X$ semidefinite representable?
- Answer: No. The feasible regions of semidefinite optimization problems are necessarily semialgebraic sets, i.e., they can be expressed as finite unions of sets defined by polynomial inequalities.
- Goal: Approximating $\log X \succeq T$ by linear matrix inequalities (LMIs)
- Want: Size of representation to grow mildly with approximation quality

Motivation

- Practice: How to solve convex optimization problems involving, e.g., quantum relative entropy?

$$D(A||B) = \mathrm{Tr}(A(\log A - \log B)) \tag{1}$$

  - No existing off-the-shelf methods
  - Semidefinite approximation of $\log X$ and related function allows solving these problems by SDP solvers, e.g., SDPT3
- Theory: Understanding approximation power of SDP
  - What can we describe with small SDPs (or SOCPs)?
  - What can we approximate with small SDPs (or SOCPs, LPs)?

Outline

Main idea

- Approximating integral representations via quadrature
  - Integral representations of functions:

$$\log x = \int_0^1 \frac{x-1}{t(x-1)+1} \mathrm{d}t$$

  - Quadrature approximation:

$$\log x \approx \sum_{j=1}^{m} w_j \frac{x-1}{t_j(x-1)+1} := r_m(x)$$

- Using functional equations to improve approximations

$$\log x^{1/2} = \frac{1}{2} \log x$$

- Extending to matrix functions and generalizing logarithm to other functions

Scalar case: approximate $\log x$

- $r_m(x)$ is monotone, concave and semidefinite representable
  - $f_t(x)$ is semidefinite representable:

$$f_t(x) := \frac{x-1}{t(x-1)+1} \geq \tau \Longleftrightarrow \begin{pmatrix} x-1-\tau & -\sqrt{t}\tau \\ -\sqrt{t}\tau & 1-t\tau \end{pmatrix} \succeq 0$$

  - $f_t(x)$ is monotone and concave
  - Gaussian quadrature: Let $t_j \in [0,1]$ be the quadrature nodes, and $w_j > 0$ be the quadrature weights.
- Exponentiation
  - when $0 < h < 1$, x^h is closer to 1 than x is, the quadrature approximation is better at x^h
  -

$$r_{m,k}(x) = 2^k r_m(\frac{x}{2^k})$$

  - $r_{m,k}(x)$ is monotone, concave and semidefinite representable

Illustration

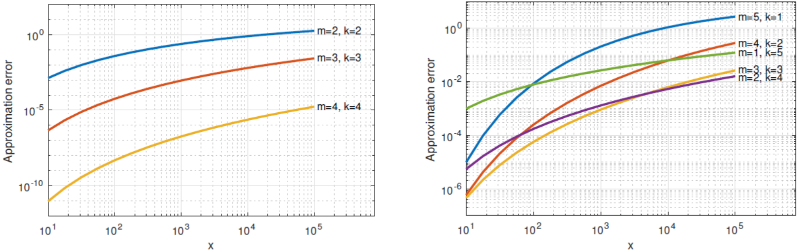

Figure 1: Plot of the error $|r_{m,k}(x) - \log(x)|$ for different choices of (m, k). Left: $m = k$. Right: pairs (m, k) such that $m + k = 6$.

Quality of approximation

Proposition 1

For any $x > 0$, we have

$$\begin{aligned} |r_{m,k}(x) - \log x| &\le 2^k |\sqrt{\kappa} - \sqrt{\kappa^{-1}}|^2 \left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1} \right)^{2m+1} \\ &\asymp 4 \cdot 4^{-m(k+2)} (\log x)^{2m+1} \end{aligned}$$

where $\kappa = x^{1/2^k}$.

Theorem 1

For any fixed $a > 1$ and $\epsilon > 0$, there exists a rational function r such that $|r(x) - \log x| < \epsilon$ for all $x \in [1/a, a]$, and r has a semidefinite representation of size $O(\sqrt{\log \frac{1}{\epsilon}})$.

- Proof. choose $k = k_1 + k_2$ where $k_1 = \log_2 \ln a + 1$, $k_2 \ge \sqrt{\log_2 \frac{32 \ln a}{\epsilon}}$, and $m = k_2/2$.

Proof of Proposition 1

- Control the error $|r(x) - \log x|$ based on the Chebyshev expansion [Trefethen, 2019]
  - Let $\log x = \int_{-1}^{1} \tilde{f}_t(x) \mathrm{d}\nu(t)$ and $r_m(x) = \int_{-1}^{1} \tilde{f}_t(x) \mathrm{d}\nu_m(t)$.
  - Compute the Chebyshev coefficients of the integrand

$$\tilde{f}_t(x) = \frac{2}{\frac{x+1}{x-1} - t} = 2\left(\sqrt{x} - 1/\sqrt{x}\right)\left(\frac{1}{2} + \sum_{k=1}^{\infty}\left(\frac{\sqrt{x}-1}{\sqrt{x}+1}\right)^k T_k(t)\right)$$

  - Let $a_k(x) = 2\left(\sqrt{x} - 1/\sqrt{x}\right)\left(\frac{\sqrt{x}-1}{\sqrt{x}+1}\right)^k$,

$$\begin{aligned}
& |\log x - r(x)| \\
= \; & \left|\sum_{k=2m}^{\infty} a_k(x)\left(\int_{-1}^{1} T_k(t)\mathrm{d}\nu(t) - \int_{-1}^{1} T_k(t)\mathrm{d}\nu_m(t)\right)\right| \\
= \; & \left|\sum_{k=m}^{\infty} a_{2k}(x)\left(\int_{-1}^{1} T_{2k}(t)\mathrm{d}\nu(t) - \int_{-1}^{1} T_{2k}(t)\mathrm{d}\nu_m(t)\right)\right| \\
\leq \; & 4|\sqrt{x} - 1/\sqrt{x}| \sum_{k=m}^{\infty}\left|\frac{\sqrt{x}-1}{\sqrt{x}+1}\right|^{2k} \\
= \; & |\sqrt{x} - 1/\sqrt{x}|^2 \left|\frac{\sqrt{x}-1}{\sqrt{x}+1}\right|^{2m-1}
\end{aligned}$$

Proof of Propsition 1 Cont.

- Scale the bound to $r_{m,k}$, let $\kappa = x^{1/2^k}$,

$$|r_{m,k}(x) - \log x| \le 2^k |\sqrt{\kappa} - 1/\sqrt{\kappa}|^2 \left|\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right|^{2m-1}$$

- Note $\left|\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right| = \left|\tanh\left(\frac{1}{4}\log\kappa\right)\right| \le \left|\frac{1}{4}\log\kappa\right| = \frac{1}{2^{k+2}}\left|\log x\right|$

$$|r_{m,k}(x) - \log x| \le |\log x| \left|\frac{\sqrt{\kappa} - 1/\sqrt{\kappa}}{2}\right|^2 \left|\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right|^{2m-1}$$

- Asymptotic error via Taylor's expansion
  - $|r_{m,k}(x) - \log x| \le$
    $2^k \left[\sinh\left(2^{-(k+1)}\log x\right)\right]^2 \left[\tanh\left(2^{-(k+2)}\log x\right)\right]^{2m-1}$
  - $$\sinh^2(2x)\tanh^{2m-1}(x) = 4x^{2m+1} + O(x^{2m+3})$$

Semidefinite approximation of relative entropy cone

- Relative entropy function: $(x, y) \in \mathbb{R}_{++} \times \mathbb{R}_{++} \mapsto x \log \frac{x}{y}$
- Relative entropy cone:

$$\mathcal{K}_{re} := \mathrm{cl}\left\{(x, y, \tau) \in \mathbb{R}^2_{++} \times \mathbb{R} : x \log \frac{x}{y} \leq \tau\right\}$$

- Let

$$\mathcal{K}_{m,k} := \mathrm{cl}\left\{(x, y, \tau) \in \mathbb{R}^2_{++} \times \mathbb{R} : x r_{m,k}(\frac{x}{y}) \leq \tau\right\}$$

Theorem 2 (Approximation error for $\mathcal{K}_{re}$)

For any fixed $a > 1$ and $\epsilon > 0$, there exist m, k with $m + k < O(\sqrt{\log \frac{1}{\epsilon}})$ such that:

1. *For all $0 < a^{-1} y < x < ay$ and $(x, y, \tau) \in \mathcal{K}_{re}$, then $(x, y, \tau + x\epsilon) \in \mathcal{K}_{m,k}$;*
2. *For all $0 < a^{-1} y < x < ay$ and $(x, y, \tau) \in \mathcal{K}_{m,k}$, then $(x, y, \tau + x\epsilon) \in \mathcal{K}_{re}$.*

Outline

Extend to matrix function

Definition 3

A matrix function g is operator concave if for any $\lambda \in [0,1]$,

$$g(\lambda A + (1-\lambda)B) \succeq \lambda g(A) + (1-\lambda) g(B)$$

and is operator monotone if $A \preceq B \Longrightarrow g(A) \preceq g(B)$

Definition 4

Given a function $g : \mathbb{R}_{++} \to \mathbb{R}$, its noncommutative (NC) perspective is $P_g : \mathbf{H}^d_{++} \times \mathbf{H}^d_{++} \to \mathbf{H}^d$ defined by

$$P_g(X,Y) = Y^{1/2} g(Y^{-1/2} X Y^{-1/2}) Y^{1/2} \tag{2}$$

- Fact 1: $\log X$ is operator concave and operator monotone
- Fact 2: If g is operator concave, then P_g is operator concave in both (X,Y). [Effros et al., 2014]

Semidefinite approximation of $\log X$

Proposition 2

Let $t \in [0,1]$ and $f_t(X) = (X - I)[t(X - I) + I]^{-1}$ is operator concave and matrix hypopgraph is semidefinite representable:

$$f_t(X) \succeq T, X \succ 0 \Longleftrightarrow \begin{bmatrix} X - I - T & -\sqrt{t}T \\ -\sqrt{t}T & I - tT \end{bmatrix} \succeq 0, X \succ 0 \quad (3)$$

- Proof. For $t \in (0,1]$, $X \succ 0$, we have

$$\begin{aligned} & f_t(X) = I/t - (I/t)[(X - I) + I/t]^{-1}(I/t) \succeq T \\ \Longleftrightarrow \quad & \begin{bmatrix} X - I + I/t & I/t \\ I/t & I/t - T \end{bmatrix} \succeq 0 \\ \Longleftrightarrow \quad & \begin{bmatrix} X - I - T & -\sqrt{t}T \\ -\sqrt{t}T & I - tT \end{bmatrix} \succeq 0 \end{aligned}$$

Semidefinite approximation of $\log X$ Cont.

- Implication: $r_m(X)$ is semidefinite representable, operator concave and operator monotone
- Generalization: For $X \succ 0, Y \succ 0$, we have

$$P_{f_t}(X,Y) \succeq T \Longleftrightarrow \begin{bmatrix} X-Y-T & -\sqrt{t}T \\ -\sqrt{t}T & Y-tT \end{bmatrix} \succeq 0$$

- Exponentiation

Proposition 3

The function $r_{m,k}$ is operator concave.

Proof.

- f_t is operator monotone and operator concave
- Power function $x \mapsto x^{1/2^k}$ is operator concave [Carlen, 2010]
- Operator concave and monotone $\circ$ Operator concave $\Longrightarrow$ Operator concave

Approximating the operator relative entropy cone

- Operator relative entropy:

$$D_{op}(X||Y) := P_{-\log}(Y, X) = -X^{1/2}\log(X^{-1/2}YX^{-1/2})X^{1/2}$$

  is jointly convex in (X, Y)
- Operator relative entropy cone:

$$K_{re}^d := \mathrm{cl}\left\{(X, Y, T) \in \mathbf{H}_{++}^d \times \mathbf{H}^d : D_{op}(X||Y) \preceq T\right\}$$

- Approximate K_{re}^d by

$$K_{m,k}^d := \mathrm{cl}\left\{(X, Y, T) \in \mathbf{H}_{++}^d \times \mathbf{H}^d : P_{-r_{m,k}}(Y, X) \preceq T\right\}$$

  where

$$P_{-r_{m,k}}(Y, X) = -X^{1/2}r_{m,k}(X^{-1/2}YX^{-1/2})X^{1/2}$$

Semidefinite approximation of K_{re}^d

Theorem 3

The cone $K_{m,k}^d$ has the following semidefinite representation:

$$(X, Y, T) \in K_{m,k}^d$$
$$\Updownarrow$$
$$\begin{array}{l} \exists T_1, ..., T_m \\ Z_1, ..., Z_k \in \mathbf{H}^d \end{array} \quad \textit{s.t.} \quad \begin{cases} Z_0 = Y \\ \begin{bmatrix} Z_i & Z_{i+1} \\ Z_{i+1} & X \end{bmatrix} \succeq 0, i = 1, ..., k-1 \\ \sum_{j=1}^{n} w_j T_j = -2^{-k} T \\ \begin{bmatrix} Z_k - X - T_j & -\sqrt{t_j} T_j \\ -\sqrt{t_j} T_j & X - t_j T_j \end{bmatrix} \succeq 0, j = 1, ..., m \end{cases} \tag{4}$$

where w_j and t_j $(j = 1, ..., m)$ are the weights and nodes for the m-point Gauss-Legendre quadrature on the interval $[0, 1]$.

Proof of Theorem 3

- Define h-weighted matrix geometric mean of $A, B \succ 0$ as

$$A \#_h B := A^{1/2}(A^{-1/2} B A^{-1/2})^h A^{1/2},$$

  i.e., the NC perspective of the power function x^h where $h \in (0,1)$.
- $A \#_h B$ is operator concave in (A, B) and semidefinite representable for rational h. [Fawzi and Saunderson, 2017]
- $$X \#_{1/2} Y \succeq T \Longleftrightarrow \exists W \in \mathbf{H}^d \text{ s.t. } \begin{bmatrix} X & W \\ W & Y \end{bmatrix} \succeq 0, W \succeq T$$

Proof of Theorem 3 Cont.

- Decomposition

$$\begin{aligned} & P_{r_{m,k}}(Y, X) \\ = \; & 2^k P_{r_m}(X \#_{1/2^k} Y, X) \\ = \; & 2^k P_{r_m}(X \#_{1/2}(X \#_{1/2} \cdots (X \#_{1/2}(X \#_{1/2} Y))), X) \end{aligned}$$

- Semidefinite representation of weighted matrix geometric means
  - $X \#_{1/2} Y$ is monotone in its second argument.

$$X \#_{2^{-k}} Y \succeq V \iff \begin{array}{c} \exists Z_0, \ldots, \\ Z_k \in \mathbf{H}^d \text{ s.t.} \end{array} \quad \begin{array}{l} \begin{bmatrix} X & Z_{i+1} \\ Z_{i+1} & Z_i \end{bmatrix} \succeq 0 \\ i = 0, \ldots k-1, \\ Z_0 = Y, Z_k \succeq V \end{array}$$

Proof of Theorem 3 Cont.

- Semidefinite representation of P_{r_m}

$$P_{r_m}(V,X) \succeq T \Longleftrightarrow \begin{array}{l} \exists T_1, ..., T_m \in \mathbf{H}^d \text{ s.t.} \\ \sum_{j=1}^{m} w_j T_j = T \\ \begin{bmatrix} V - X - T_j & -\sqrt{t_j} T_j \\ -\sqrt{t_j} T_j & X - t_j T_j \end{bmatrix} \succeq 0 \\ j = 1, ...m \end{array}$$

- P_{r_m} is monotone in its first argument.

Semidefinite approximation of quatum relative entropy

- Recall $D(A||B) = \mathrm{Tr}(A(\log A - \log B))$

Proposition 4 ([Tropp, 2015])

Given $A, B \in \mathbf{H}^d_{++}$, *we have* $D(A||B) = \phi(D_{op}(A \otimes I || I \otimes \bar{B}))$ *where* ϕ *is the unique linear map from* $\mathbb{C}^{d^2 \times d^2} \to \mathbb{C}$ *such that* $\phi(X \otimes Y) = \mathrm{Tr}(XY^T)$, $\bar{B}$ *is the elementwise complex conjugate of* B.

Corollary 1

$$D(A||B) \le \tau$$
$$\Updownarrow$$
$$\exists T \in \mathbf{H}^{d^2} \text{ s.t. } (A \otimes I, I \otimes \bar{B}, T) \in K^{d^2}_{re}, \phi(T) \le \tau$$

- Fact: $X \preceq Y \implies \phi(X) \le \phi(Y)$

Beyond matrix logarithm

- It is possible to extend such approximation idea to other functions
- Recall two pillars of approximation:
  - Quadrature approximation of integral representation
  - Functional equation
- Any operator monotone function admits an integral representation in terms of rational functions

Theorem 4 (Löwner's Theorem [Hansen and Pedersen, 1982])

If $g : \mathbb{R}_{++} \longrightarrow \mathbb{R}$ is a operator monotone function, then there is a unique probability measure ν supported on $[0,1]$ such that $g(x) = g(1) + g'(1) \int_0^1 f_t(x) \mathrm{d}\nu(t)$.

- It is also possible to find other functional equations like $P_g \circ \Phi = P_g$ where g is positive operator monotone and Φ has contraction and monotonicity properties [Fawzi et al., 2019].

Outline

Maximum entropy problems

$$\begin{array}{rll} \max & -\sum_{i=1}^{n} x_i \log x_i & \\ \text{s.t.} & Ax = b & (A \in \mathbb{R}^{\ell \times d}) \\ & x \geq 0 & \end{array}$$

- $m = k = 3$

		Successive approximation (CVX)		Padé approximation (this paper)		
n	ℓ	time (s)	accuracy	time (s)	accuracy	$\lvert p_{sa} - p_{Pade} \rvert$
50	25	0.34 s	1.065e-05	0.32 s	1.719e-06	8.934e-06
100	50	0.52 s	1.398e-06	0.34 s	2.621e-06	1.222e-06
200	100	1.10 s	6.635e-06	0.88 s	2.767e-06	3.868e-06
400	200	3.38 s	2.662e-05	0.72 s	1.164e-05	1.498e-05
600	300	9.14 s	2.927e-05	1.84 s	2.743e-05	1.843e-06
1000	500	52.40 s	1.067e-05	3.91 s	1.469e-04	1.362e-04

- CVX's successive approximation : use Taylor expansion

Geometric programming

$$
\begin{array}{lll}
\min & \sum_{k=1}^{w_0} c_{0,k} x^{a_0,k} & \\
\text{s.t.} & \sum_{k=1}^{w_j} c_{j,k} x^{a_j,k} \leq 1 & (j = 1, \dots, \ell) \\
& x \geq 0 &
\end{array}
$$

			Successive approximation (CVX)		Padé approximation (this paper)		
n	ℓ	sp	time (s)	accuracy	time (s)	accuracy	$\lvert p_{sa} - p_{Pade} \rvert$
50	50	0.3	1.28 s	2.509e-07	0.94 s	2.106e-06	1.856e-06
50	100	0.3	1.78 s	2.045e-05	1.03 s	3.122e-05	1.077e-05
100	100	0.1	1.57 s	4.759e-06	1.16 s	5.197e-06	4.383e-07
100	150	0.1	3.60 s	8.484e-06	1.60 s	2.240e-06	6.244e-06
100	200	0.1	7.60 s	1.853e-06	2.69 s	3.769e-06	1.916e-06
200	200	0.1	7.47 s	2.441e-07	3.72 s	7.505e-07	9.945e-07
200	400	0.1	42.71 s	3.666e-06	14.36 s	2.855e-06	6.521e-06
200	600	0.1	184.33 s	7.899e-06	35.45 s	4.480e-06	3.419e-06

Table 3: Geometric programming (34) using our method (with $(m, k) = (3, 3)$) and the successive approximation scheme of CVX, on different random instances. The column "sp" indicates the sparsity of the power vectors $a_{j,k}$ (i.e., how many variables appear in each monomial terms). Also we used $w_0 = w_1 = \cdots = w_\ell = 5$ (i.e., the posynomial objective as well as the posynomial constraints all have 5 terms). Accuracy is measured via absolute error between the optimal value returned by the approximation and the built-in MOSEK solver for geometric programs (`mskgpopt`).

Variational formula for trace

$$\mathrm{Tr}(Y) = \max_{X \succ 0} \mathrm{Tr}(X) - D(X||Y)$$

```
cvx_begin
  variable X(n,n) symmetric
  maximize (trace(X) - quantum_rel_entr(X,Y))
cvx_end
```

n	time (s)	accuracy
5	2.37 s	1.143e-06
10	4.32 s	2.844e-06
15	9.56 s	4.732e-06
20	24.39 s	7.537e-06
25	77.02 s	9.195e-06
30	163.07 s	1.290e-05

Table 4: Result of solving the optimization problem (35) for different Hermitian positive definite matrices Y of size $n \times n$ with $\mathrm{Tr}[Y] = 1$. The problems were implemented using CVX as shown above and solved using SDPT3. The accuracy column reports the quantity $|p - 1|$ where p is the optimal value returned by the solver (note that the matrix Y is sampled to have trace one).

available at https://github.com/hfawzi/cvxquad

Further topics

- Specific topics:
  - Lower bounds: Given $\epsilon > 0$, what is the smallest $s(\epsilon)$ such that there exists SOCP representable f with $\max_{x\in[1/e,e]} |f(x) - \log x| < \epsilon$?
  - Self-concordant barriers for cone K_{re}^d?
  - Smaller semidefinite approximations for $D(X||Y)$?
- Broad issues:
  - What can we describe with small SDPs (or SOCPs)?
  - What can we approximate with small SDPs (or SOCPs)?
  - How to approximate and preserve structural properties?

Conclusion

- Principled approximations scheme
- Complexity of SDP approximation grows mildly with approximation quality
- Matrix logarithm has ϵ-approximate semidefinite description with $O(\sqrt{\log \frac{1}{\epsilon}})$ $2n \times 2n$ LMIs.
- A new SOCP approximation for relative entropy cone

Thank You! Questions?

Further Reading I

Ben-Tal, A. and Nemirovski, A. (2001).
Lectures on modern convex optimization: analysis, algorithms, and engineering applications.
SIAM.

Carlen, E. (2010).
Trace inequalities and quantum entropy: an introductory course.
Entropy and the quantum, 529:73–140.

Effros, E., Hansen, F., et al. (2014).
Non-commutative perspectives.
Ann. Funct. Anal, 5(2):74–79.

Fawzi, H. and Saunderson, J. (2017).
Lieb's concavity theorem, matrix geometric means, and semidefinite optimization.
Linear Algebra and its Applications, 513:240–263.

Further Reading II

Fawzi, H., Saunderson, J., and Parrilo, P. A. (2019).
Semidefinite approximations of the matrix logarithm.
Foundations of Computational Mathematics, 19(2):259–296.

Hansen, F. and Pedersen, G. K. (1982).
Jensen's inequality for operators and löwner's theorem.
Mathematische Annalen, 258(3):229–241.

Trefethen, L. N. (2019).
Approximation theory and approximation practice, volume 164.
Siam.

Tropp, J. A. (2015).
An introduction to matrix concentration inequalities.
arXiv preprint arXiv:1501.01571.