Lecture 9: Quantum entropy maximization
CSE 599S: Entropy optimality, Winter 2016
Instructor: James R. Lee **Last updated:** February 15, 2016

1 Quantum states and von Neumann entropy

Recall that $\mathcal{S}_{\mathrm{sym}} \subseteq \mathbb{R}^{n\times n}$ is the set of real, symmetric $n \times n$ matrices. Let $\mathrm{spec}(A)$ denote the set of eigenvalues of A. Recall that $\mathcal{S}^n_{++} \subseteq \mathcal{S}^n_{+}$ denote the subsets of positive definite and positive semi-definite matrices, respectively. We write $A \succ 0$ and $A \succeq 0$ to denote the respective memberships $A \in \mathcal{S}^n_{++}$ and $A \in \mathcal{S}^n_{+}$.

A matrix $Q \in \mathcal{S}^n_{+}$ with $\mathrm{Tr}(Q) = 1$ is called a *density matrix* and it is the basic object of quantum information theory. Note that if Q were diagonal, it would correspond in the natural way to a classical probability distribution over n objects. A general density matrix Q represents the quantum state of a system with n degrees of freedom.

Quantum measurements. Although we will not need this, it may help to think about a quantum state operationally. A *resolution of the identity* if a decomposition $\mathrm{Id} = \sum_{i=1}^{k} \rho_i$, where $\rho_1, \ldots, \rho_k \succeq 0$. One can perform a "quantum measurement" corresponding to such a resolution, where the probability of obtaining outcome i is $\mathrm{Tr}(\rho_i Q)$.

1.1 Trace convexity

In order to define the quantum notions of entropy and relative entropy, it helps to understand that convex functionals on $\mathbb{R}_+$ can be lifted to convex functionals on $\mathcal{S}^n_{+}$.

For an interval $J \subseteq \mathbb{R}$, define the set $\mathcal{S}_{\mathrm{sym}}(J) = \{A \in \mathcal{S}_{\mathrm{sym}} : \mathrm{spec}(A) \in J\}$. Now any function $f : J \to \mathbb{R}$ can be defined on $\mathcal{S}_{\mathrm{sym}}(J)$ as follows. If $A = P^T D P$, where $D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is a diagonal matrix containing the eigenvalues of A, then

$$f(A) \stackrel{\mathrm{def}}{=} P^T \begin{pmatrix} f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_n) \end{pmatrix} P\,.$$

This allows us to define the *von Neumann entropy* of a density Q as

$$\mathcal{S}(Q) = -\mathrm{Tr}(Q \log Q) = -\sum_i \lambda_i \log \lambda_i\,,$$

where $\{\lambda_i\}$ denote the eigenvalues of Q, and we continue to use the convention that $0 \log 0 = 0$. Again, observe that for diagonal matrices, this is just the usual Shannon entropy. In a moment, we will prove the following.

Lemma 1.1. *For an interval $J \subseteq \mathbb{R}$, if the function $f : J \to \mathbb{R}$ is (strictly) convex and continuous, then the function $A \mapsto \mathrm{Tr}[f(A)]$ is (strictly) convex and continuous on $\mathcal{S}_{\mathrm{sym}}(J)$.*

This will imply that the function $Q \mapsto \mathrm{Tr}(Q \log Q)$ is strictly convex on $\mathcal{S}_+^n$. In particular, it gives rise to the Bregman divergence

$$\begin{aligned}\mathcal{S}(A \,\|\, B) &= \mathrm{Tr}(A \log A) - \mathrm{Tr}(B \log B) - \mathrm{Tr}\left((A-B)\nabla \mathrm{Tr}(B \log B)\right) \\ &= \mathrm{Tr}(A(\log A - \log B)) - \mathrm{Tr}(A-B)\,.\end{aligned}$$

If $\mathrm{Tr}(A) = \mathrm{Tr}(B) = 1$, this yields the *quantum relative entropy between A and B:*

$$S(A \,\|\, B) \stackrel{\mathrm{def}}{=} \mathrm{Tr}(A(\log A - \log B))\,.$$

In particular, by general properties of Bregman divergences, we have $S(A \,\|\, B) \geqslant 0$ and $S(A \,\|\, B) = 0 \implies A = B$.

Proof of Lemma 1.1. First of all, since f is continuous, the map $A \mapsto \mathrm{Tr}[f(A)]$ is continuous, and we need only verify that it is midpoint convex (since then general convexity follows by continuity).

Let $\{u_j\}$ be an orthornormal basis of eigenvectors for $(A+B)/2$. We will use the convexity of f twice. First:

$$\begin{aligned}\mathrm{Tr}\left[f\left(\frac{A+B}{2}\right)\right] &= \sum_{j=1}^{n} f\left(\left\langle u_j, \frac{A+B}{2} u_j\right\rangle\right) \\ &\leqslant \frac{1}{2}\sum_{j=1}^{n} f\left(\langle u_j, A u_j\rangle\right) + f\left(\langle u_j, B u_j\rangle\right)\,.\end{aligned}$$

The next calculation holds for *any* orthonormal basis $\{u_j\}$. Let $\{v_i\}$ be an orthonormal eigenbasis for A and write $u_j = \sum_i \langle u_j, v_i\rangle v_i$. Use convexity of f again to write:

$$\begin{aligned}f\left(\langle u_j, A u_j\rangle\right) &= f\left(\sum_{i=1}^{n} \langle u_j, v_i\rangle^2 \langle v_i, A v_i\rangle\right) \\ &\leqslant \sum_{i=1}^{n} \langle u_j, v_i\rangle^2 f\left(\langle v_i, A v_i\rangle\right)\,,\end{aligned}$$

where we have used the fact that $\sum_i \langle u_j, v_i\rangle^2 = \|u_j\|^2 = 1$. Now sum both sides over j and use $\sum_j \langle u_j, v_i\rangle^2 = \|v_i\|^2 = 1$ to conclude that

$$\sum_{j=1}^{n} f\left(\langle u_j, A u_j\rangle\right) \leqslant \sum_{i=1}^{n} f\left(\langle v_i, A v_i\rangle\right) = \mathrm{Tr}[f(A)]\,.$$

Applying the same reasoning for B completes the proof. One can easily observe from the argument that strict convexity of f yields strict convexity of $\mathrm{Tr}[f(A)]$ as well. □

2 Quantum state approximation and entropy maximization

We will use $\mathcal{U} = \mathrm{Id} \,/\, \mathrm{Tr}(\mathrm{Id})$ to denote the "maximally mixed state." By minimizing $\mathcal{S}(P \,\|\, \mathcal{U})$ over all density matrices P subject to linear constraints, one can prove the following, in analogy with our approximation of high-entropy distributions.

Theorem 2.1. *Suppose that* $Q \in \mathcal{S}_+^n$ *satisfies* $\mathrm{Tr}(Q) = 1$, *and let* $\{F_1, F_2, \ldots\} \subseteq \mathcal{S}_{\mathrm{sym}}^n$ *denote a sequence of tests. Given any* $\varepsilon > 0$, *consider the quantum entropy optimization problem:*

$$\begin{aligned} \textit{minimize} \quad S(P \,\|\, U) \quad \textit{subject to} \quad & P \succeq 0, \mathrm{Tr}(P) = 1 \\ & \mathrm{Tr}(F_i P) \leqslant \mathrm{Tr}(F_i Q) + \varepsilon \quad \forall i \,. \end{aligned}$$

The unique optimal solution satisfies

$$P^* = \frac{\exp\left(-\sum_i c_i F_i\right)}{\mathrm{Tr}\left[\exp\left(-\sum_i c_i F_i\right)\right]} \,,$$

where $\{c_i \geqslant 0\}$ *are non-negative constants satisfying*

$$\sum_i c_i \leqslant \frac{S(Q \,\|\, \mathcal{U})}{\varepsilon} \,.$$

This can be proved relatively easily using duality theory for convex programs (and, in particular, Slater's condition, where the presence of Q and $\varepsilon > 0$ implies that strong duality holds). Already the key calculation occurs in the case of a single test; let us prove this special case.

Proof. Assume there is a single test F_1. We claim that the unique optimal solution to the optimization problem

$$\text{minimize} \quad \mathcal{S}(P \,\|\, \mathcal{U}) + \mathrm{Tr}(FP) \quad \text{subject to } P \succeq 0, \mathrm{Tr}(P) = 1$$

is given by $P^* = e^{-F} / \mathrm{Tr}(e^{-F})$. If we prove this, then we can set $F = cF_1$ where $c = \mathcal{S}(P \,\|\, \mathcal{U})/\varepsilon$.

By minimality, we conclude that

$$\mathcal{S}(P^* \,\|\, \mathcal{U}) + c\, \mathrm{Tr}(F_1 P^*) \leqslant \mathcal{S}(Q \,\|\, \mathcal{U}) + c\, \mathrm{Tr}(F_1 Q) \,,$$

which implies, since $\mathcal{S}(P^* \,\|\, \mathcal{U}) \geqslant 0$, that

$$\mathrm{Tr}(F_1 P^*) \leqslant \mathrm{Tr}(F_1 Q) + \varepsilon \,,$$

completing the argument.

So let us now verify that P^* is the minimizer of $\Psi(P) := \mathrm{Tr}(P \log P) + \mathrm{Tr}(FP)$ over density matrices. The original objective function above only differs from this one by an additive constant. First, evaluate $\Psi(P^*) = -\log \mathrm{Tr}(e^{-F})$.

Then observe that since quantum relative entropy is always non-negative, we have

$$0 \leqslant S(P \,\|\, P^*) = \mathrm{Tr}(P \log P) + \mathrm{Tr}(PF) + \log \mathrm{Tr}(e^{-F}) \,,$$

which is equivalent to

$$\Psi(P) \geqslant -\log \mathrm{Tr}(e^{-F}) = \Psi(P^*) \,,$$

showing that indeed P^* is minimal. □

2.1 Approximation by a low-degree square

Let us now use Theorem 2.1 to approximate a function $Q : \{0,1\}^n \to \mathcal{S}_+^r$ by a low-degree square. Recall that for a function $F : \{0,1\}^n \to \mathcal{S}_{\mathrm{sym}}^r$, we denote $\deg(F(x)) = \max_{ij} \deg(F(x)_{ij})$.

Theorem 2.2. *Suppose that $Q : \{0,1\}^n \to \mathcal{S}_+^r$ satisfies $\mathbb{E}_x \operatorname{Tr}(Q(x)) = 1$, and furthermore we are given tests $F_1, F_2, \ldots : \{0,1\}^n \to \mathcal{S}_{\mathrm{sym}}^r$. Then for every $\varepsilon > 0$, there exists a function $R : \{0,1\}^n \to \mathcal{S}_+^r$ such that $\mathbb{E}_x \operatorname{Tr}(R(x)^2) = 1$ and for all $i = 1, 2, \ldots$, we have*

$$\mathop{\mathbb{E}}_x \operatorname{Tr}(F_i(x)R(x)^2) \leqslant \mathop{\mathbb{E}}_x \operatorname{Tr}(F_i(x)Q(x)) + \varepsilon \,,$$

and furthermore

$$\deg(R) \leqslant O(1)\frac{1 + \mathbb{E}_x \operatorname{Tr}(Q(x))S(Q(x) \,\|\, \mathcal{U})}{\varepsilon} \cdot \max_{x,i} \|F_i(x)\| \cdot \max_i \deg(F_i)\,. \tag{2.1}$$

Note that here $R(x)^2$ refers to the matrix product square. This is the analog of the junta-approximation theorem we saw for high-entropy distributions. To prove it using Theorem 2.1, one needs to convert Q and $\{F_1, F_2, \ldots\}$ to block-diagonal matrices as follows:

$$\bar{Q} = \mathop{\mathbb{E}}_x Q(x) \otimes e_x e_x^T$$
$$\bar{F}_i = \sum_x F_i(x) \otimes e_x e_x^T$$

Applying Theorem 2.1 with approximation parameter $\varepsilon/2$ (and then unpacking the block-diagonal optimizer), one gets an approximator of the form

$$P^*(x) = \frac{\exp\left(-\sum_i c_i F_i(x)\right)}{\operatorname{Tr}\left[\exp\left(-\sum_i c_i F_i(x)\right)\right]}\,,$$

with $\sum_i c_i \leqslant 2\frac{S(\bar{Q} \,\|\, \mathcal{U})}{\varepsilon}$, and $\mathcal{S}(\bar{Q} \,\|\, \mathcal{U}) = \mathbb{E}_x \operatorname{Tr}(Q(x))\mathcal{S}(Q(x) \,\|\, \mathcal{U})$.

Now we approximate $P^*(x)$ by a low-degree square by approximating $e^x = (e^{x/2})^2$ by a truncated Taylor expansion of $e^{x/2}$. The degree of truncation is determined by the approximation parameter $\varepsilon/2$ and the maximum eigenvalue of the exponent, hence the form of the bound (2.1).

3 Hyperbolic cones and factorization scaling

We have examined two positive cones so far: The positive orthant $\mathbb{R}_+$ and the PSD cone $\mathcal{S}_+^n$. One might ask if there are other natural cones over which we might do entropy optimization.

A prime example is the set of *hyperbolic cones.* For definitions and properties of these cones, please refer to the links on the web page. For now, let us simply mention how they relate to the present setting.

Consider a degree-d homogeneous polynomial $p \in \mathbb{R}[x_1, \ldots, x_n]$. Say that p is *hyperbolic in direction* $e \in \mathbb{R}^n$ if the univariate polynomial $t \mapsto p(x - et)$ has all real roots for every $x \in \mathbb{R}^n$. In this case, let $\lambda_1(x) \leqslant \lambda_2(x) \cdots \leqslant \lambda_d(x)$ denote these roots, and define the set

$$\Lambda_+(e) = \{x \in \mathbb{R}^n : \lambda_1(x) \geqslant 0\}\,.$$

It turns out that $\Lambda_+(e)$ is a closed convex cone (called a hyperbolic cone).

It carries a natural entropy functional $H_\Lambda(x) = -\sum_{i=1}^n \lambda_i(x) \log \lambda_i(x)$. And this is a concave function on $\Lambda_+(e)$. [Again, see the links on the web page until formal references are added here.]

Given a cone $\mathcal{K} \subseteq \mathbb{R}^n$, we recall its *dual cone* $\mathcal{K}^* = \{x \in \mathbb{R}^n : \langle x, y \rangle \geqslant 0 \ \forall y \in \mathbb{R}^n\}$. One can speak in general of a matrix $M : X \times Y \to \mathbb{R}_+$ *factoring through* $\mathcal{K}$ as follows:

$$M(x, y) = \langle u_x, v_y \rangle$$

for some vectors $\{u_x\} \subseteq \mathcal{K}$ and $\{v_y\} \subseteq \mathcal{K}^*$. A natural question to ask is whether such a factorization can be rescaled to be analytically "nice" when $\mathcal{K}$ is a hyperbolic cone. In the following section, we describe a rescaling for factorizations through the PSD cone that is a crucial preprocessing step before applying Theorem 2.2 to prove a lower bound on $\bar{\gamma}_{\mathrm{sdp}}$ for the cut polytope.

It is an interesting open question whether there is analogous rescaling for hyperbolic cones.

Question 3.1 (Hyperbolic cone rescaling)**.** Let $\mathcal{K} \subseteq \mathbb{R}^n$ be a hyperbolic cone of degree d, and suppose $\{u_1, \ldots, u_s\} \subseteq \mathcal{K}$ and $\{v_1, \ldots, v_t\} \subseteq \mathcal{K}^*$ are such that $\langle u_i, v_j \rangle \geqslant 0$ for all $i \in [s], j \in [t]$. Is it the case that there exists *a possibly different* degree-d hyperbolic cone $\hat{\mathcal{K}} \subseteq \mathbb{R}^n$ and vectors $\{\hat{u}_1, \ldots, \hat{u}_s\} \subseteq \hat{\mathcal{K}}$ and $\{\hat{v}_1, \ldots, \hat{v}_t\} \subseteq \hat{\mathcal{K}}^*$ such that $\langle u_i, v_j \rangle = \langle \hat{u}_i, \hat{v}_j \rangle$ for all $i \in [s], j \in [t]$, and furthermore

$$\max_{i,j} \|\hat{u}_i\| \cdot \|\hat{v}_j\| \leqslant \mathrm{poly}(n, d) \cdot \max_{i,j} \langle u_i, v_j \rangle \quad ?$$

This is already interesting when the defining hyperbolic polynomial p is *multi-linear*, in which case one can assume that $d \leqslant n$.

3.1 John's theorem and factorization rescaling

Finite-dimensional operator norms. Let H denote a finite-dimensional Euclidean space over $\mathbb{R}$ equipped with inner product $\langle \cdot, \cdot \rangle$ and norm $|\cdot|$. For a linear operator $A : H \to H$, we define the operator, trace, and Frobenius norms by

$$\|A\| = \max_{x \neq 0} \frac{|Ax|}{x}, \quad \|A\|_* = \mathrm{Tr}(\sqrt{A^T A}), \quad \|A\|_F = \sqrt{\mathrm{Tr}(A^T A)}\,.$$

Let $\mathcal{M}(H)$ denote the set of self-adjoint linear operators on H. Note that for $A \in \mathcal{M}(H)$, the preceding three norms are precisely the ℓ_∞, ℓ_1, and ℓ_2 norms of the eigenvalues of A. For $A, B \in \mathcal{M}(H)$, we use $A \succeq 0$ to denote that A is positive semi-definite and $A \succeq B$ for $A - B \succeq 0$. We use $\mathcal{D}(H) \subseteq \mathcal{M}(H)$ for the set of density operators: Those $A \in \mathcal{M}(H)$ with $A \succeq 0$ and $\mathrm{Tr}(A) = 1$.

One should recall that $\mathrm{Tr}(A^T B)$ is an inner product on the space of linear operators, and we have the operator analogs of the Hölder inequalities: $\mathrm{Tr}(A^T B) \leqslant \|A\| \cdot \|B\|_*$ and $\mathrm{Tr}(A^T B) \leqslant \|A\|_F \|B\|_F$.

Rescaling PSD factorizations. As in the case of non-negative rank, consider finite sets X and Y and a matrix $M : X \times Y \to \mathbb{R}_+$. For the purposes of proving a lower bound on the psd rank of some matrix, we would like to have a nice analytic description.

To that end, suppose we have a rank-r psd factorization

$$M(x, y) = \mathrm{Tr}(A(x)B(y))$$

where $A : X \to \mathcal{S}_+^r$ and $B : Y \to \mathcal{S}_+^r$. The following result of Briët, Dadush and Pokutta (2013) gives us a way to "scale" the factorization so that it becomes nicer analytically. (The improved bound stated here is from an article of Fawzi, Gouveia, Parrilo, Robinson, and Thomas, and we follow their proof.)

Lemma 3.2. *Every M with $\mathrm{rank}_{\mathsf{psd}}(M) \leqslant r$ admits a factorization $M(x,y) = \mathrm{Tr}(P(x)Q(y))$ where $P : X \to \mathcal{S}_+^r$ and $Q : Y \to \mathcal{S}_+^r$ and, moreover,*

$$\max\{\|P(x)\| \cdot \|Q(y)\| : x \in X, y \in Y\} \leqslant r\|M\|_\infty ,$$

where $\|M\|_\infty = \max_{x\in X, y\in Y} M(x,y)$.

Proof. Start with a rank-r psd factorization $M(x,y) = \mathrm{Tr}(A(x)B(y))$. Observe that there is a degree of freedom here, because for any invertible operator J, we get another psd factorization $M(x,y) = \mathrm{Tr}\left((JA(x)J^T) \cdot ((J^{-1})^T B(y)J^{-1})\right)$.

Let $U = \{u \in \mathbb{R}^r : \exists x \in X \; A(x) \succeq uu^T\}$ and $V = \{v \in \mathbb{R}^r : \exists y \in X \; B(y) \succeq vv^T\}$. Set $\Delta = \|M\|_\infty$. We may assume that U and V both span $\mathbb{R}^r$ (else we can obtain a lower-rank psd factorization). Both sets are bounded by finiteness of X and Y.

Let $C = \mathrm{conv}(U)$ and note that C is centrally symmetric and contains the origin. Now John's theorem tells us there exists a linear operator $J : \mathbb{R}^r \to \mathbb{R}^r$ such that

$$B_{\ell_2} \subseteq JC \subseteq \sqrt{r} B_{\ell_2} , \tag{3.1}$$

where B_{ℓ_2} denotes the unit ball in the Euclidean norm. Let us now set $P(x) = JA(x)J^T$ and $Q(y) = (J^{-1})^T B(y) J^{-1}$.

Eigenvalues of $P(x)$:. Let w be an eigenvector of $P(x)$ normalized so the corresponding eigenvalue is $\|w\|_2^2$. Then $P(x) \succeq ww^T$, implying that $J^{-1}w \in U$ (here we use that $A \succeq 0 \implies SAS^T \succeq 0$ for any S). Since $w = J(J^{-1}w)$, (3.1) implies that $\|w\|_2 \leqslant \sqrt{r}$. We conclude that every eigenvalue of $P(x)$ is at most r.

Eigenvalues of $Q(y)$:. Let w be an eigenvector of $Q(y)$ normalized so that the corresponding eigenvalue is $\|w\|_2^2$. Then as before, we have $Q(y) \succeq ww^T$ and this implies $J^T w \in V$. Now, on the one hand we have

$$\max_{z \in JC} \langle z, w\rangle \geqslant \|w\|_2 \tag{3.2}$$

since $JC \supseteq B_{\ell_2}$.

On the other hand:

$$\max_{z\in JC} \langle z, w\rangle^2 = \max_{z\in C} \langle Jz, w\rangle^2 = \max_{z\in C} \langle z, J^T w\rangle^2 . \tag{3.3}$$

Finally, observe that for any $u \in U$ and $v \in V$, we have

$$\langle u, v\rangle^2 = \langle uu^T, vv^T\rangle \leqslant \max_{x\in X, y\in Y} \langle A(x), B(y)\rangle \leqslant \Delta .$$

By convexity, this implies that $\max_{z\in C} \langle z, v\rangle^2 \leqslant \Delta$ for all $v \in V$, bounding the right-hand side of (3.3) by Δ. Combining this with (3.2) yields $\|w\|_2^2 \leqslant \Delta$. We conclude that all the eigenvalues of $Q(y)$ are at most Δ. ☐