Freie Universität Berlin
Tutorials on Quantum Information Theory
Winter term 2020/21

Problem Sheet 4
Kraus representation and Norms for matrices part I

J. Eisert, J. Haferkamp, J. C. Magdalena De La Fuente

1. **On the Kraus representation of quantum channels**
The operational meaning of Kraus operators can be understood in the following setting in which, for simplicity, we restrict ourselves to quantum channels with the same input and output space $L(\mathcal{X})$. Suppose we apply a unitary U to the joint system and environment in the state $\rho \otimes |0\rangle\langle 0| \in L(\mathcal{X} \otimes \mathcal{Z})$, where $|0\rangle \in \mathcal{Z}$ is some reference state, and then we measure system $\mathcal{Z}$ in the computational basis.

a) Show that the action of the unitary on the joint system can be written as

$$U(\rho \otimes |0\rangle\langle 0|)U^\dagger = \sum_{kl} E_k \rho E_l^\dagger \otimes |k\rangle \langle l| \ ,$$

with respect to the basis $\{|i\rangle\}_i$ on the second system.

Solution: We define $E_k = (\mathbb{1} \otimes \langle k|)U(\mathbb{1} \otimes |0\rangle)$ and check.

b) Now, we perform a von-Neumann measurement on $\mathcal{Z}$ in the same basis. Determine the post-measurement state conditioned on outcome i. What is the probability of obtaining outcome i?

Solution: Up to normalisation the post-measurement state is given by $\rho_i = E_k \rho E_k^\dagger$. The probability reads $p(i|\rho) = \mathrm{Tr}[(\mathbb{1} \otimes |i\rangle\langle i|)U(\rho \otimes |0\rangle\langle 0|)U^\dagger = \mathrm{Tr}[E_i^\dagger E_i \rho]$

c) Give the corresponding operational interpretation of the Kraus operators E_k and the unitary U.

Solution: The $E_k^\dagger E_k$ can be seen as elements of a POVM implemented on the first system by the von-Neumann measurement on the second system.

d) Now, suppose we want to implement a von-Neumann measurement on $\mathcal{X}$ via a global unitary and a von-Neumann measurement on $\mathcal{Z}$. Characterize the unitaries $U \in U(\mathcal{X} \otimes \mathcal{Z})$ on the joint system that give rise to this situation. Give an example for the case of two qubits.

Solution: They have to satisfy

$$[(\mathbb{1} \otimes \langle i|)U(\mathbb{1} \otimes |0\rangle)]^2 = (\mathbb{1} \otimes \langle i|)U(\mathbb{1} \otimes |0\rangle),$$
$$[(\mathbb{1} \otimes \langle i|)U(\mathbb{1} \otimes |0\rangle)]^\dagger = (\mathbb{1} \otimes \langle i|)U(\mathbb{1} \otimes |0\rangle).$$

An example is $CX = |0\rangle\langle 0| \otimes \mathbb{1} + |1\rangle\langle 1| \otimes X$.

Finally, we will show some properties of the Kraus representation

e) Let $\{K_i\}_{i=1}^N$ and $\{\tilde{K}_j\}_{i=1}^N$ be two sets of linear operators in $L(\mathcal{X}, \mathcal{Z})$ fulfilling the completeness relation of Kraus operators. Show that if the two sets are related by a unitary transformations $U \in U(N)$ such that $\tilde{K}_i = \sum_j U_{ij} K_j$, the channels represented by the sets coincide.

Solution: We have

$$\begin{aligned}\tilde{T}(X) &= \sum_j \tilde{K}_j X \tilde{K}_j^\dagger = \sum_{ijk} U_{ij} K_j X K_k^\dagger \overline{U}_{ik} \\ &= \sum_{jk} \left(U_{ki}^\dagger U_{ij} \right) K_j X K_k^\dagger = \sum_{jk} \delta_{kj} K_j X K_k^\dagger,\end{aligned}$$

where in the last equality we used the unitarity of U.

f) Show that all equal-sized Kraus representations of a given channel T are related via a unitary transformation.

Hint: Relate the Kraus representation of two low-rank matrix factorisations of the Choi matrix.

Solution: Let $\{K_i\}$ and $\{\tilde{K}_i\}$ be sets of Kraus operators. The Choi matrices of the corresponding quantum channel Φ can be expressed as

$$J(\Phi) = AA^\dagger = \tilde{A}\tilde{A}^\dagger \tag{1}$$

where the matrix A is given by

$$A = \big(\operatorname{vec} K_1, \operatorname{vec} K_2, \dots, \operatorname{vec} K_N\big) \tag{2}$$

and $\tilde{A}$ defined analogously with the $\tilde{K}$s.

Recall from linear algebra that two low-rank factorisations $L_1R_1 = L_2R_2$ of the same matrix are always related by a invertible matrix G such that $L_1 = L_2G$ and $R_1 = G^{-1}R_2$. For the case Hermitian matrices with $R_i = L_i^\dagger$, we conclude that G must be a unitary matrix.

Thus, there exists a unitary matrix $U(N)$ such that $\tilde{A} = AU$.

2. **ℓ_p-norms**

In quantum information we deal with a handful of different matrix spaces such as the set of quantum states and also quantum channels. For quantitative statements we have to equip these spaces with distance measures. Depending on the application and context different distance measures have the desired operational meaning.

A prominent role is played by the so called *Schatten p-norms*. But to set the stage we have to first familiarise ourself with their analogons on vector spaces, namely *ℓ_p-norms*. For $1 \leq p < \infty$ the ℓ_p-norm on the complex vector space $\mathbb{C}^n$ is defined as

$$||\bullet||_{\ell_p} : x \mapsto ||\bullet||_{\ell_p} := \left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}},$$

and the ℓ_∞-norm as

$$||\bullet||_{\ell_\infty} : x \mapsto ||\bullet||_{\ell_\infty} := \lim_{p\to\infty} ||\bullet||_{\ell_p}.$$

We will now characterise the function $||\bullet||_{\ell_p}$ and derive important properties. We begin with an explicit expression for the ℓ_∞-norm.

a) Show that $||\bullet||_{\ell_\infty} = \max_{1\leq i\leq n} |x_i|$.

Solution: We assume w.l.o.g. that $|x_1| = \max_i |x_i|$

$$||x||_{\ell_\infty} = \lim_{p\to\infty} \left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}} \tag{3}$$

$$= |x_1| \lim_{p\to\infty} \left[1 + \sum_{i=2}^n \frac{|x_i|^p}{|x_1|^p}\right]^{\frac{1}{p}} \tag{4}$$

$$= |x_1|. \tag{5}$$

For all of what follows the notion of a convex function will be important. Let $D \subset \mathbb{R}$ be a convex set. We say that a function $f : D \to \mathbb{R}$ is *convex* if

$$f\left(\sum_i a_i x_i\right) \leq \sum_i a_i f(x_i),$$

for all $x_i \in D$ and $a_i \geq 0, i = 1, \ldots, m$ such that $\sum_i a_i = 1$.

b) Show that any twice continuously differentiable function on an open interval is convex if and only if its second derivative is everywhere nonnegative.

Solution: First, observe that the definition above is equivalent to requiring that for $\lambda \in (0, 1)$ and $x, y \in D$ it holds that $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$. The definition follows directly by applying this result repeatedly.

Convex $\Leftarrow f'' \geq 0$

For any $x, y \in (\alpha, \beta)$, $x < y$, and $\lambda \in (0, 1)$, we set $z = \lambda x + (1 - \lambda)y$. Assume $f''(x)$ be non-negative on (α, β), correspondingly $f'(x)$ is non-decreasing on (α, β). Now,

$$\begin{aligned} f(z) &= \lambda f(z) + (1 - \lambda)f(z) && (6)\\ &= \lambda \int_x^z f'(t)\,\mathrm{d}t + \lambda f(x) + (1 - \lambda)\int_y^z f'(t)\,\mathrm{d}t + (1 - \lambda)f(y) && (7)\\ &\leq \lambda f(x) + (1 - \lambda)f(y) + \lambda f'(z)(z - x) + (1 - \lambda)f'(z)(z - y) && (8)\\ &= \lambda f(x) + (1 - \lambda)f(y) + f'(z)\left[z - (\lambda x + (1 - \lambda)y)\right] && (9)\\ &= \lambda f(x) + (1 - \lambda)f(y). && (10) \end{aligned}$$

Convex $\Rightarrow f'' \geq 0$

Conversely, assume that $f''(x)$ is negative somewhere, than by continuity there exist a subinterval (α', β') where f' is decreasing everywhere. Choosing $x, y \in (\alpha', \beta')$ and $\lambda \in (0, 1)$ implies that $\int_x^z f'(t)\,\mathrm{d}t > f'(z)(z - x)$ and $\int_y^z f'(t)\,\mathrm{d}t > f'(z)(z - y)$. Thus, the same calculation as above yields $f(z) > \lambda f(x) + (1 - \lambda)f(y)$ establishing a contradiction to f being convex.

Alternatively: Assume $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$ holds. Then, the following is also true

$$\begin{aligned} f(x) &= f\left(\frac{1}{2}(x + h) + \frac{1}{2}(x - h)\right) && (11)\\ &\leq \frac{1}{2}f(x + h) + \frac{1}{2}f(x - h) && (12)\\ \Leftrightarrow 0 &\leq f(x + h) + f(x - h) - 2f(x). && (13) \end{aligned}$$

This is exactly the term one encounters in the Taylor expansion of the second derivative

$$f''(x) = \lim_{h \to \infty} \frac{f(x + h) + f(x - h) + 2f(x)}{h^2}. \quad (14)$$

c) Show that $|\bullet|^p$ is a convex function for $p \geq 1$.

Solution: Applying the criterion of the last excercise, it is obvious that the function $(0, \infty) \to \mathbb{R}$ $x \mapsto x^p$ with $p \geq 1$ is convex. Consider $x, y \in \mathbb{R}$, $x, y \neq 0$ and $\lambda \in (0, 1)$ then

$$|\lambda x + (1-\lambda)y|^p \leq (\lambda|x| + (1-\lambda)|y|)^p \tag{15}$$
$$\leq \lambda|x|^p + (1-\lambda)|y|^p. \tag{16}$$

We will now use this fact to show that $|| \bullet ||_{\ell_p}$ is a norm (positive definite, absolutely homogeneous, subadditive aka triangle inequality).

d) Argue that $|| \bullet ||_{\ell_p}$ is positive definite and absolutely homogeneous for $1 \leq p < \infty$ and $p = \infty$.

That was easy. Now comes the hard part; we have to show that the norms satisfy the triangle inequality, i.e.

$$||x + y||_{\ell_p} \leq ||x||_{\ell_p} + ||y||_{\ell_p}. \tag{17}$$

In fact, the triangle inequality for ℓ_p-norms has even its own name, *Minkowski inequality*. A clever way to prove this inequality is to normalise the right hand side, introduce normalised vectors and then use the convexity of $|\cdot|^p$.

e) Argue that it is sufficient to consider the case $||x||_{\ell_p} = \lambda$ and $||y||_{\ell_p} = (1-\lambda)$ with $\lambda \in (0, 1)$ in order to prove the Minkowski inequality.

Solution: Let $\tilde{x}, \tilde{y} \in \mathbb{R}$ then by absolute homogeneity of $||x||_{\ell_p}$ we have

$$||\tilde{x} + \tilde{y}||_{\ell_p} \leq ||\tilde{x}||_{\ell_p} + ||\tilde{y}||_{\ell_p} \tag{18}$$

if and only if

$$||\frac{s}{\tilde{x}} + \frac{s}{\tilde{y}}||_{\ell_p} \leq 1 \tag{19}$$

with $s := ||\tilde{x}||_{\ell_p} + ||\tilde{y}||_{\ell_q}$. Choosing $x = \frac{s}{\tilde{x}}$ and $y = \frac{s}{\tilde{y}}$ yields the situation above.

f) Show the Minkowski inequality for the ℓ_p-norms when $1 \leq p < \infty$.

Solution: We write $x = \lambda\hat{x}$ and $y = (1-\lambda)\hat{y}$ with $||\hat{x}||_{\ell_p} = ||\hat{y}||_{\ell_p} = 1$. Then,

$$||x + y||_{\ell_p}^p = ||\lambda\hat{x} + (1-\lambda)\hat{y}||_{\ell_p}^p \tag{20}$$
$$= \sum_i |\lambda\hat{x}_i + (1-\lambda)\hat{y}_i|^p \tag{21}$$
$$\leq \sum_i [\lambda|\hat{x}_i|^p + (1-\lambda)|\hat{y}_i|^p] \tag{22}$$
$$= \lambda||\hat{x}||_{\ell_p} + (1-\lambda)||\hat{y}||_{\ell_p} \tag{23}$$
$$= \lambda + (1-\lambda) = 1. \tag{24}$$

A crucial property of the ℓ_p-norms is Hölder's inequality. It generalises the Cauchy-Schwarz inequality, which is its special case for $p = 2$. Let $\langle \cdot, \cdot \rangle$ be the Euclidean inner product on $\mathbb{C}^n$, i.e. $\langle x, y \rangle = \sum_{i=1}^n \bar{x}_i y_i$ with $\bar{\cdot}$ denoting the complex conjugate. Hölder's inequality reads

$$|\langle x, y \rangle| \leq ||x||_{\ell_p} ||y||_{\ell_q}, \text{ where } \frac{1}{p} + \frac{1}{q} = 1.$$

Like for the proof of Minkowski's inequality, it will be useful to use normalised vectors in the proof of Hölder's inequality. Furthermore, we will need to first establish the *arithmetic-geometric mean inequality*

$$\prod_{i=1}^{n} x_i^{a_i} \leq \sum_{i=1}^{n} a_i x_i \text{ if } x_i \geq 0,\ a_i \geq 0,\ \sum_i a_i = 1. \tag{25}$$

g) Show that $-\log$ is a convex function and use this to show the arithmetic-geometric mean inequality, Eq. (25).

Solution: The function $-\log$ is twice continously differential and $(-\log x)'' = (-1/x)' = 1/x^2 > 0$ for $x \in \mathbb{R}$ and, thus, convex.

Then,

$$-\log\left[\prod_i x_i^{a_i}\right] = -\sum_i a_i \log x_i \geq -\log\left[\sum_i a_i x_i\right]. \tag{26}$$

By monotonicity of the logarithm, this implies Eq. (25).

h) Now, prove Hölder's inequality for $1 < p < \infty$.

Solution: Again, by absolute homogeneity of the norms and bilinearity of the scalar product, it is sufficient to consider the case $\|x\|_{\ell_p} = \|y\|_{\ell_q} = 1$.

$$|\langle x, y\rangle| = \sum_i |x_i||y_i| = \sum_i (|x_i|^p)^{1/p}(|y_i|^q)^{1/q}) \tag{27}$$

$$\leq \sum_i (\frac{1}{p}|x_i|^p + \frac{1}{q}|y_i|^q) = \frac{1}{p}||x||_{\ell_p}^p + \frac{1}{q}||y||_{\ell q}^q \tag{28}$$

$$= \frac{1}{p} + \frac{1}{q} = 1. \tag{29}$$

i) Finally, prove Hölder's inequality for $p = 1$.

Solution: $|\langle x, y\rangle| = \sum_i |x_i||y_i| \leq \max_i\{|x_i|\} \sum_i |y_i| = ||x||_{\ell_\infty} ||y||_{\ell_1}$.

More generally, for a norm $\|\cdot\|$ on $\mathbb{C}^d$ one can define its dual norm $\|\cdot\|^*$ as

$$\|x\|^* := \sup_{y \in \mathbb{C}^d, \|y\|=1} |\langle x, y\rangle|. \tag{30}$$

j) Show that for every norm $\|\cdot\|$ on $\mathbb{C}^d$ it holds:

$$|\langle x, y\rangle| \leq \|x\|\|y\|^* \tag{31}$$

for all $x, y \in \mathbb{C}^d$.

k) Show that the dual norm $|| \bullet ||_{\ell_p}^*$ of the ℓ_p-norm $||\bullet_{\ell_p}$ is the ℓ_q-norm $||\bullet_{\ell_q}$ with $\frac{1}{p} + \frac{1}{q} = 1$.

Solution: By Hölder's inequality, we have

$$\|x\|_{\ell_q} \geq \sup_{\|y\|_{\ell_p}=1} |\langle x, y\rangle|. \tag{32}$$

Again by absolute homogeneity we can assume $\|x\|_{\ell_q} = 1$. Setting $y_i = |x_i|^{q/p+1}/x_i$ for all nonzero x_i and 0 else one checks that for $1/p + 1/q = 1$ the inequality is saturated establishing the claim:

$$|\langle x, y\rangle| = \left|\sum_i x_i \frac{|x_i|^{q/p+1}}{x_i}\right| = \sum_i |x_i|^{q/p+1} = \sum_i |x_i|^q = ||x||_{\ell_q}^q = 1$$

Finally, we will show another convenient property of the ℓ_p norms.

l) Show that the ℓ_p norms are ordered in the sense that

$$||x||_{\ell_p} \leq ||x||_{\ell_q}, \text{ for } q \leq p.$$

Solution: Consider $x \in \mathbb{C}^n$ and define $\hat{x} = x/||x||_{\ell_q}$. In particular, we have $\hat{x}_i \leq 1$ for all i. Then, $||x||_{\ell_p}^p = ||x||_{\ell_q}^p \sum_i |\hat{x}_i|^p \leq ||x||_{\ell_q}^p \sum_i |\hat{x}_i|^q = ||x||_{\ell_q}^p ||\hat{x}||_{\ell_q}^q = ||x||_{\ell_q}^p$.