6.443/8.371/18.436 Quantum Information Science II Lecture 3

Classical and quantum error correction

Aram Harrow February 14, 2018

3.1 Norms in QC

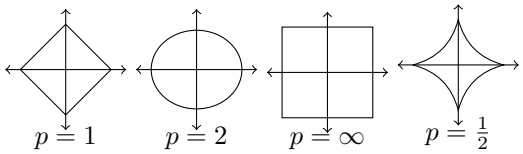
Before we start error correction it is useful to go over norms. This sections is not covered in the textbook, but it provides some perspective for types of operations and objects. It can explain the similar conditions governing measurements and density matrices. We will look at the norms of these objects.

Norm is a function, $\| \star \|$, that:

1. Is homogenous: $\|cv\| = \|c\| \times \|v\|$
2. Obeys the triangle inequality: $\|v + w\| \leq \|v\| + \|w\|$
3. Separating: $\|v\| = 0$ implies $v = 0$.

It provides the way of measuring a size of an object.

An example of norm functions on $\mathbb{C}^d$ are the l_p norms: $l_p = (\sum_i |v_i|^p)^{1/p}$. The l_2 norm is the familiar Euclidean distance. The l_1 norm is the sum of absolute values, sometimes referred to as the Manhattan distance. As $p \to \infty$, the largest elements are given more and more weight, so l_∞ returns the max element. The value of p determines how balanced the measurement is: whether weight is given to large or small elements. The unit circle corresponding to each of these norms is according to the following



The l_p norms are defined on vectors, so what about norms for matrices? The Schatten-P norms, S_p. Given a matrix X the singular value decomposition of X is $X = U\Sigma V^\dagger$, where U and V are isometries, and $\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_n)$ for non-negative numbers σ_i. To calculated $\|X\|_{S_p}$ for a matrix X, take the singular value decomposition, and take the l_p norm of those: $\|X\|_{S_p} = \|\mathrm{sv}(X)\|_{l_p} = (\sum_{i=1}^n \sigma_i^p)^{1/p}$. For hermitian X, we can just talk about the eigenvalues of X. If $X \succeq 0$ then $\|X\|_{S_1} = \mathrm{tr}X$.

This gives another way of describing some of the objects we've seen so far:

- **Density Matrices**: Positive semidefinite matrices on the unit sphere of S_1 (Schatten-1 norm is just trace).
$$\rho \text{ is a density matrix } \iff \rho \succeq 0 \text{ and } \|\rho\|_{S_1} = 1$$
- **Measurement Operators**: The intersection of positive semidefinite matrices and the unit ball of S_∞.
$$M \text{ is a measurement operator } \iff 0 \preceq M \preceq I \iff M \succeq 0 \text{ and } \|M\|_{S_\infty} \leq 1$$
We often write $\mathrm{tr}(M\rho) = \langle M, \rho \rangle$.

3.1.1 Duality

Given norm $\|\|$ we can define its corresponding dual norm according to

$$\|x\|_* = \max_{\|y\| \leq 1} | \langle x, y \rangle |$$

This means $| \langle x, y \rangle | \leq \|x\| \cdot \|y\|_*$. l_2 is self-dual. l_1 and l_∞ and S_1 and S_∞ are dual to each other.

3.1.2 Trace distance

The following defines the trace distance

Definition 1 (Trace distance)**.** The trace distance of two density matrices ρ, σ is defined as

$$T(\rho, \sigma) := \frac{1}{2}\|\rho - \sigma\| = \frac{1}{2}\mathrm{tr}|\rho - \sigma|, \tag{3.1}$$

where $\|\cdot\|$ is the Schatten norm for $p = 1$. Here for a matrix A, $|A| = \sqrt{AA^\dagger}$.

This is useful because for any measurement M,

$$\begin{aligned} |\operatorname{tr}(M\rho) - \operatorname{tr}(M\sigma)| = |\operatorname{tr}(M(\rho - \sigma))| & \\ &\leq \|M\|_{S_\infty}\|\rho - \sigma\|_{S_1} \\ &\leq \|\rho - \sigma\|_{S_1} \\ &= 2T(\rho, \sigma) \end{aligned}$$

The factor 2 in the above is not needed. As a matter of fact you will prove in the problem set that

$$T(\rho, \sigma) = \max_M |\operatorname{tr}(M(\rho - \sigma))|.$$

The trace distance is contractive under quantum operations N:

$$T(N(\rho), N(\sigma)) \leq T(\rho, \sigma). \tag{3.2}$$

Any unitary operation will preserve the trace distance, while for example, the depolarizing operator will shorten the trace distance. In this sense, it is inevitable that the information will be lost when we apply quantum operations. Error correction can slow down this process.

3.2 Classical Error Correction

The objective is to encode string of bits with larger strings of bits in a way that we can correct the effect of noise. In general noise can be either random (e.g. flip each bit with independent probability p) or worst case (any error up to $\leq l$ positions). An encoding map is a map like $E : \{0,1\}^k \to \{0,1\}^n$. A decoding map $D : \{0,1\}^n \to \{0,1\}^k$ is designed to correct most or all errors. We may sometimes demand that encoding and decoding maps are efficient to implement. But we should think of this as an extra feature.

One example is the repetition code. Consider for example $k = 1$:

- Encode: $E(0) = 0^n$ and $E(1) = 1^n$
- Decode: $D(x) = \mathrm{MAJORITY}(x)$ (Majority vote e.g. $110 \to 1$)

There are two ways to characterize the performance of this code:

- (Deterministic) This code corrects any one error (n-bit code corrects $\frac{n-1}{2}$ error).
- (Random noise) Suppose there are independent and identically distributed (IID) bit flips (on physical bits) with probability p.

  We can analyze the performance of this code

$$\begin{aligned} \Pr[D(\mathrm{noise}(E(bit))) \neq \mathrm{bit}] &:= \Pr[\mathrm{error}] && (3.3)\\ &:= \Pr[\mathrm{Bin}(n,p) \geq n/2] && (3.4)\\ &\leq \exp\bigl(-n/2(1/2-p)^2\bigr) && (3.5) \end{aligned}$$

  in the last line we have used the Chernoff bound.

In this next several lectures, we will focus on the deterministic approach, but when we discuss information theory later, we will discuss the case of random errors. Also here we drop the efficiency requirement for now.

From the intuition from the previous example, we give the definition of classical codes.

Definition 2 (Classical code)**.** A $[n, k, d]$ classical code C is a set that satisfies

- Codewords: $C = \mathrm{Im}(E) \subseteq \{0,1\}^n$, which means C uses n physical bits.
- $|C| = 2^k$, which means C encodes k logical bits.
- Minimum distance $d := \min_{\substack{x \neq y \\ x,y \in C}} \mathrm{dist}(x, y)$, where $\mathrm{dist}(x, y) = \|x - y\|_1$ is the Hamming distance[1].

Claim 3. *A $[n, k, d]$ code corrects any $\frac{d-1}{2}$ error, and detects any $d - 1$ error.*

Proof. For detection, suppose we start with $x \in C$ and apply error e to get $\tilde{X} = X + e \mod 2$. Our objective is to detect whether $e = 0$ or not. We want to show that if $\|e\|_1 \leq d - 1$ then we can detect it. It is just enough to test if $\tilde{X} \in C$. We detect that an error has occurred if $\tilde{X} \notin C$. However if $\|e\|_1 \geq d$ this scheme doesn't work, because the error can result in $\tilde{X} \in C$.

For correction, we use the decoding map that outputs the nearest codeword. Let us analyze this scheme. Suppose we start with $X \in C$ and then apply noise e with $\|e\|_1 \leq \frac{d-1}{2}$ and get $\tilde{X}$. Then $\mathrm{dist}(X, \tilde{X}) \leq \frac{d-1}{2}$. Our decoder outputs $y = \arg\min_{Z \in C} \mathrm{dist}(Z, \tilde{X})$. Now suppose as a way of contradiction that $Y \neq X$. Therefore

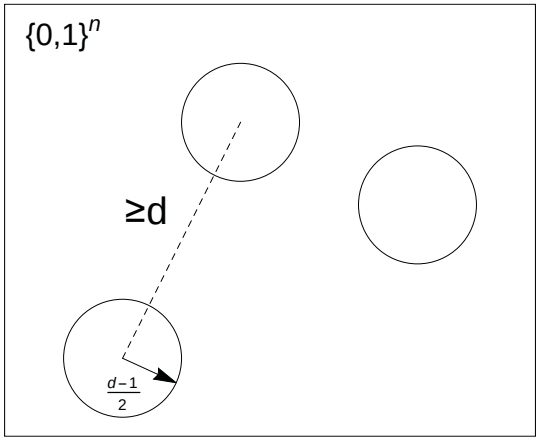
$$d \leq \mathrm{dist}(X, Y) \leq \mathrm{dist}(X, \tilde{X}) + \mathrm{dist}(Y, \tilde{X}) \leq \frac{d-1}{2} + \mathrm{dist}(Y, \tilde{X})$$

hence

$$\mathrm{dist}(Y, \tilde{X}) \geq \frac{d+1}{2}.$$

which is a contradiction. Conversely $d/2$ cannot be corrected.

It is helpful to think about this proof geometrically. The centers of the balls are the codewords in C. Strings within the disjoint balls can be corrected because there is a unique codeword within distance $\frac{d-1}{2}$.



□

The proof suggests the relation between the high-dimensional sphere packing problem with the error correction.

[1]The Hamming distance is the number of positions at which the corresponding symbols are different.

3.3 Linear codes

All the arithmetic operations in this section are done on finite field $\mathbb{F}_2 = \{0, 1\}$. $\mathbb{F}_2$ means that arithmetic is done modulo 2, i.e. after each operation we divide by 2 and take the remainder, which is always 0 or 1. Addition and multiplication is almost the same as normal over the integers except that $1 + 1 = 0$. Note also that $x = -x$. Note that unlike $\{0, 1\}$, $\mathbb{F}_2$ is closed under addition and multiplication. Moreover, unlike $\mathbb{Z}$, $\mathbb{F}_2$ has multiplication inverses. Therefore $\mathbb{F}_2$ is a field, as a result linear algebra works.

A linear code C is a subspace $C \leq \mathbb{F}_2^n$ and $\dim C = k$. $C = \text{span}\{g_1, g_2, \cdots, g_k\}$, where $g_1, \cdots, g_k$ are generators of the code. $x \in C$ is equivalent to $x = a_1 g_1 + \cdots + a_k g_k$, where $a_1, \cdots, a_k \in \mathbb{F}_2$. $|C| = 2^k$.

Example 1. $C = \{000, 111\} \subseteq \mathbb{F}_2^3$. The generator of the code is $G = (1, 1, 1)^T$. $G[0] = (0, 0, 0)^T$ and $G[1] = (1, 1, 1)^T$.

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in C \Leftrightarrow \begin{cases} x_1 = x_2 \\ x_2 = x_3 \end{cases} \Leftrightarrow \begin{cases} x_1 + x_2 = 0 \\ x_2 + x_3 = 0 \end{cases} \Leftrightarrow \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0.$$

$H = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$ is the parity check matrix and the code C is the Kernel of H (i.e. the set of x satisfying $Hx = 0$).

The above example suggests that we can define the code in two ways. The generator matrix is good for encoding ($x \in \mathbb{F}_2^k$ is encoded as Gx) and the parity check matrix is good for error-detection, since it gives a list of parities that should be zero for valid codewords. Using this approach $C = \{Ga : a \in \mathbb{F}_2^k\} = \text{Im}\, G$, where $G = (g_1 | \ldots | g_k)$. Therefore $G \in \mathbb{F}_2^{n \times k}$. Another way to describe the code is by the matrix H satisfying $Hx = 0$ for all $x \in C$. In other words $C = \text{Ker}(H)$. $H \in \mathbb{F}_2^{n \times (n-k)}$