Statistical Inference in Quantum Computing

Alistair Wallis

February 1, 2010

Motivation

- Traditional computing: Information sent as bits - 0's and 1's.
- Quantum computing: Information sent as qubits - these can be 0, 1 or a superposition of 0 and 1
- Some believe these ideas could lead to very fast computers
- When measuring system, we have the system in a state with a probability
- Makes sense therefore to consider questions relating to probabilities and statistics

State vectors

Postulate 1: With every isolated physical system we associate an abstract Hilbert space $\mathcal{H}$ over the field $\mathbb{C}$. The state of the system is exhaustively characterised by a normalised state vector $|\psi\rangle$. These state vectors follow the superposition principle.

Superposition Principle: If $|\psi_1\rangle, \ldots, |\psi_n\rangle$ are possible state vectors, then any normalised linear combination of these vectors is again a state vector. For example,
$|\psi\rangle = \frac{1}{\sqrt{n}}|\psi_1\rangle + \frac{1}{\sqrt{n}}|\psi_2\rangle + \ldots + \frac{1}{\sqrt{n}}|\psi_n\rangle$ is a normalised linear combination of these vectors, and therefore this describes a new state vector.

Qubits

The simplest quantum system is a *qubit* (QUantum BIT) (also known as the *spin-half* system). A qubit has a two dimensional state space, with orthonormal basis $\{|0\rangle, |1\rangle\}$. An arbitrary state vector can therefore be written as

$$|\psi\rangle = a|0\rangle + b|1\rangle,$$

where $a, b \in \mathbb{C}$ and $|a|^2 + |b|^2 = 1$, the normalisation condition. If one imagines that each $|\psi\rangle$ is an element of $\mathbb{C}^n$, one sees that the normalisation condition is equivalent to $\langle\psi|\psi\rangle = 1$.

Density operator

Now let us suppose our finite-dimensional quantum system is prepared in one of a number of states $\{|\psi_k\rangle\}$, each with a probability of p_k, with the probabilities p_k summing to one. We call $\{p_k, |\psi_k\rangle\}$ an *ensemble of pure states* and define the following matrix, ρ, called the *density matrix*, with corresponding operator called the *density operator*.

$$\rho := \sum_k p_k |\psi_k\rangle\langle\psi_k|.$$

A quantum system whose state $|\psi\rangle$ is known exactly is said to be in a *pure state*. In this case, $\rho = |\psi\rangle\langle\psi|$. Otherwise, ρ is in a *mixed state* and is said to be a *mixture* of the different pure states in the ensemble for ρ.

Density Operator vs State Vector

Theorem

An operator ρ is the density operator associated to an ensemble $\{p_k, |\psi_k\rangle\}$ if and only if it satisfies the following conditions:

1. $\text{tr}(\rho) = 1$
2. ρ is positive semi-definite

Alternative Postulate 1

Postulate 1′: With every isolated physical system we associate an abstract Hilbert space $\mathcal{H}$ over the field $\mathbb{C}$. The system is completely described by its density operator ρ, which is a positive semi-definite, self-adjoint operator with trace 1 acting on the state space of the system. If the quantum system is in state ρ_j with probability p_j, then the density operator associated with the system is

$$\rho = \sum_j p_j \rho_j .$$

Postulate 2

Postulate 2: The evolution of a closed quantum system is described by a unitary transformation. That is, the state $|\psi_1\rangle$ of the system at time t_1 is related to the state $|\psi_2\rangle$ of the system at time t_2 by a unitary operator U which depends only on the times t_1 and t_2,

$$|\psi_2\rangle = U|\psi_1\rangle.$$

Postulate 2′: The time evolution of a closed quantum system is described by the Schrödinger equation,

$$i\hbar \frac{d|\psi\rangle}{dt} = H|\psi\rangle.$$

Alternative Postulate 2

Postulate 2$''$: The evolution of a closed quantum system is described by a unitary transformation. That is, the state ρ_1 of the system at time t_1 is related to the state ρ_2 of the system at time t_2 by a unitary operator U which depends only on the times t_1 and t_2,

$$\rho_2 = U \rho_1 U^\dagger.$$

Postulate 3

Postulate 3: The state space of a composite physical system is the tensor product of the state spaces of the component physical systems. Moreover, if we have systems numbered 1 through n, and system number k is prepared in the state ρ_k, the joint state of the total system is $\rho_1 \otimes \rho_2 \otimes \ldots \otimes \rho_n$.

Entanglement

Suppose we have the following state

$$|\psi\rangle = \frac{|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle}{\sqrt{2}}$$

This state has the property that there are no single qubit states $|a\rangle$ and $|b\rangle$ such that $|\psi\rangle = |a\rangle \otimes |b\rangle$

Definition

A state of a composite system with the property that it cannot be written as a product of states of its component systems is called an *entangled state*.

Operator-valued probability measure

Definition

Let (Ω, Σ) be a measurable space. An *operator-valued probability measure* (oprom) (a.k.a generalised measurement, positive operator valued measure) $\mathcal{M}$ is a collection of self-adjoint matrices $\mathcal{M}(\Gamma)$ $(\Gamma \in \Sigma)$ such that the following hold

1. $\mathcal{M}(\Omega) = I$
2. $\mathcal{M}(\Gamma)$ is positive semi-definite for all $\Gamma \in \Sigma$
3. For a sequence $(\Gamma_n)_n$ of disjoint elements of Σ,

$$\mathcal{M}\left(\bigcup_n \Gamma_n\right) = \sum_n \mathcal{M}(\Gamma_n).$$

Postulate 4

Postulate 4: A measurement $\mathcal{M}$ on a d-dimensional quantum system taking values from a measurable space (Ω, Σ) is described by an operator-valued probability measure. If a measurement is performed, then the probability distribution of the outcome X is given by

$$\mathcal{P}(X \in \Gamma) = \mathrm{tr}(\rho \mathcal{M}(\Gamma)),$$

where $\Gamma \in \Sigma$.

Theorem

The probability distribution described in postulate 4 is a probability measure on the measurable space (Ω, Σ).

Likelihood

Definition

We define the *likelihood function* for the single variable X to be

$$L(\theta; x) := f(x; \theta).$$

Definition

We define the *loglikelihood function* for the single variable X to be

$$l(\theta; x) := \log(L(\theta; x)).$$

Score and Information

Definition

We define the *score* $V(\theta, X)$ for the single variable X to be

$$V(\theta, X) := l'(\theta; X).$$

Setting the score to zero gives a stationary point of $l(\theta; x)$ and this gives us one way of finding an estimate $\hat{\theta}$ of θ.

Definition

We define the *expected Fisher information* $I_X(\theta)$ *in the random variable* X to be

$$I_X(\theta) := \mathrm{Var}(V(\theta; x)).$$

Cramér-Rao inequality

Theorem

Let $(x_1, \dots, x_n)$ be a sample of n independent observations of a random variable X whose probability function at x is $f(x, \theta)$, where θ is an unknown parameter. Let $T(x_1, \dots, x_n)$ be an unbiased estimator of θ. Then, subject to a set of regularity conditions about the integrals and differentials of f,

$$\mathrm{Var}(T) \geq \frac{1}{nI(\theta)}.$$

Likelihood

Suppose $\rho = \rho(\theta)$, where θ is some unknown parameter. Suppose $\mathcal{M}$ is an oprom taking values from a measurable space (Ω, Σ) describing a measurement on our quantum system. So, by the fourth postulate, we have

$$\mathcal{P}(X \in \Gamma) = \mathrm{tr}(\rho(\theta)\mathcal{M}(\Gamma)),$$

where $\Gamma \in \Sigma$. We will define μ by

$$\mu(\Gamma) = \mathrm{tr}(\mathcal{M}(\Gamma))$$

and assume that $M(\Gamma)$ can be written as

$$M(\Gamma) = \int_{\Gamma} m(x) d\mu(x),$$

where m is a positive semi-definite, self-adjoint operator.

We then have that ρ has probability density with respect to μ given by

$$f(x;\theta) = \mathrm{tr}(\rho(\theta)m(x)).$$

Assuming the relevant integrals and derivatives exist, the expected Fisher information will be

$$I_M(\theta) = \mathbb{E}((V(\theta;X))^2) = \int_\Omega (V(\theta;X))^2 f(x;\theta)d\mu(x),$$

where $l(\theta;X) := \log(f(X;\theta))$ is the loglikelihood and $V(\theta;X) := \frac{\partial}{\partial\theta}(l(\theta;X))$ is the score function.

Quantum Analogues

Definition

We define the *quantum score function* (or *symmetric logarithmic derivative*) λ of ρ with respect to θ to be the self-adjoint operator which is the solution to the equation

$$\dot{\rho} = \frac{\partial \rho(\theta)}{\partial \theta} = \frac{1}{2}(\rho\lambda + \lambda\rho).$$

Definition

The *expected quantum information number* $I_Q(\theta)$ (also known as the *Helstrom information number*) is defined by

$$I_Q(\theta) := \text{tr}(\rho\lambda^2).$$

Big Theorem

Theorem

$I_M(\theta) \leq I_Q(\theta)$

Theorem

If T is an unbiased estimator of θ, then

$$\mathrm{Var}(T) \geq \frac{1}{I_Q(\theta)}.$$